Design of an Echo Canceller For Acoustic Halls

by

Ilke Levent Akin

A Thesis Presented to the

FACULTY OF THE COLLEGE OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

SYSTEM ENGINEERING

April, 1993

Order Number 1354115

Design of an echo canceller for acoustic halls

Akin, Ilke Levent, M.S.

King Fahd University of Petroleum and Minerals (Saudi Arabia), 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

DESIGN OF AN ECHO CANCELLER FOR ACOUSTIC HALLS

BY

ILKE LEVENT AKIN

A Thesis Presented to the

FACULTY OF THE COLLEGE OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

In

SYSTEMS ENGINEERING

APRIL, 1993

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

This thesis, written by

İlke Levent AKIN

under the direction of his Thesis Advisor and approved by his Thesis Committee, has been presented to and accepted by the Dean of the College of Graduate Studies, in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN SYSTEMS ENGINEERING

Thesis Committee

Dr. M.A. El – Shafei (Chairman)

Dr. D. Kavranoğlu (Member)

Dr. M.S. El – Hennawey (Member)

Department Chairman

Dean, College of Graduate Studies

Date : 1-5-93

Dedicated to

The Great TURKISH Nation

Acknowledgment

I would like to express special thanks to my excellent thesis committee chairman Dr. M. A. El-Shafei for his patient guidance, constructive criticism, and generous support during the preparation of this thesis.

I would also like to thank my thesis committee members Dr. D. Kavranoğlu and Dr. M. S. El-Hennawey for their support.

I wish to thank the Systems Engineering Department members and students for the cooperation they have shown and who provided support throughout my studies. I wish to thank the Electrical Engineering Department for allowing me to use the department's facilities. Acknowledgment is due to King Fahd University of Petroleum and Minerals for support of this research.

I am also thankful to my friends in Saudi Arabia who made my long stay pleasant and memorable.

Contents

List of Tables

List of Figures

23

Abstract

Name : İlke Levent AKIN

Title : Design of an Echo Canceller for Acoustic Halls

Major Field : Systems Engineering

Date of Degree : April 1993

A major problem in many public address systems of large halls is the poor perception of message due to multiple reflections - **echoes** - in the hall. The general approach to this problem is to identify the characteristics of the echo path and then generate an echo replica which cancels out the real echo signal.

In this thesis, a novel structured FIR equalizer is proposed for the cancellation process. It requires less taps than the existing equalizers and operates directly on the dominant echoes to generate an efficient echo replica. Extensive simulation studies are carried out for evaluating the performance of the proposed acoustic echo canceller. The new approach considerably improves the computational complexity and enables possible implementations on commercial DSP boards.

Master of Science Degree

King Fahd University of Petroleum and Minerals

Dhahran, Saudi Arabia

April 1993

خلاصة الرسالة

اسم الطالب : اليكي لفنت اكيني
عنوان الرسالة : تصميم جهاز لإزالة صدى الصوت فى القاعات الواسعة .
التخصص : هندسة النظم .
تاريخ الشهادة : أبريل ١٩٩٣ م .

من المشاكل الرئيسية فى نظم الصوتيات بالاساكن العامة صعوبة تمييز الرسائل الصوتيه نتيجة لتعدد الانعكاسات وهو ما يعرف بصدى الصوت . وتعتمد الطرق الحديثه لمعالجة هذه المشكله بصورة عامة على التعرف على خواص المسارات الصوتيه ثم تصنيع صدى مشابه تماما ومعاكس لإزالة الصدى الحقيقى .

وفى هذه الرسالة نعرض تركيبا جديدا لمعادل صوتى باستعمال مرشح ذات استجابه قصيرة لإزالة الصوت . وهذا المرشح يحتاج لعدد أقل من المعاملات بالمقارنه بالمرشحات التى اقترحت سابقا . كما إنه يتعامل مباشرة مع موجات الصدى الرئيسية فقط لعمل صدى صوتى معاكس بكفاءة عاليه . وقد تضمنت هذه الرسالة أيضا دراسة وافيه بالمحاكاة التقنين أداء المعادل الصوتي . وقد دلت هذه الدارسة على أن الطريقه المقترحة تقلل من التعقيد الحسابى مما يمكن من تنفيذها بسهولة باستعمال لوحات معالجة الاشارات الرقميه المتوفرة .

درجة الماجستير في العلوم
جامعة الملك فهد للبترول والمعادن
الظهران ، المملكة العربية السعودية
أبريل ١٩٩٣ م

Chapter 1

INTRODUCTION

1.1 Introduction

A major problem in many public address systems of large halls, like airports, train stations, closed stadiums, etc., is the poor perception of message due to multiple reflections in the hall. Techniques to improve the perception of speech in such difficult acoustic environments have therefore been attracting the researcher's interest since two decades. Thus, a new technique is proposed for the improvement of perception in the acoustic halls of public address systems. The technique is designed to achieve acoustical echo equalization which renders itself for implementation using commercial Digital Signal Processors (DSPs), and it will also be capable of reshaping the overall transfer function of the acoustic hall to approximate certain desirable acoustic characteristics.

1.2 Background and Problem Formulation

The study of acoustics for sound systems is classified into outdoors and indoors with indoors divided into large-room and small-room acoustics. In acoustics, reflection of a delayed and distorted version of an original sound or signal to the listener is known as the *echo phenomenon.* This phenomenon exists in the public telephone networks and in the public address systems. However, the physical properties of echoes in telephone and acoustical networks differs mainly from the features of echoes, because of the different nature of the echo paths.

Acoustical echo equalization is a research field that follows a rapid growth. Efforts to set up real-time and economic acoustical echo canceller have been increasing in the literature, since the last two decades. Equalization of the acoustical echoes of a large hall is a difficult and complex problem.

A typical acoustic echo canceller requires a large number of computations in a very limited time with a high accuracy. The techniques to cancel the acoustical echoes are in the field of adaptive filtering technology and they require efficient adaptive filters that can converge fast. In order to model the acoustic path of a large hall, the adaptive filter generally requires a large length of a few thousands. For instance, in a usual acoustical hall where the sampling rate is 8 kHz., it is required to have an adaptive filter of a few thousands taps which are adaptively controlled for maximum echo cancellation [1,2]. This kind of filter makes the adaptation process impractical for the computational complexity and for the proportionality between the number of taps and the convergence time. Adaptively controlling such a large number of taps requires complex algorithms and also for a better quality of perception this filter requires larger number of taps which increases the convergence time.

In addition to the length of adaptive filter, a typical impulse response of an acoustical echo path makes it obvious that the acoustic echo cancellation is a challenging task. The duration of the impulse response of the acoustic echo path is usually in the range of 100 to 400 msec. Moreover, the response may change rapidly, for instance due to a door opening or a person moving, at any time. In order to get even a reasonable improvement, it is required to have an adaptive filter with a few thousands of taps [1].

As for now most widely used techniques in the acoustical echo cancellation use the adaptive Finite Impulse Response (FIR) filter for the modelling, this is also referred as a tapped delay line. However, one of the major disadvantages of using such a filter is that as the delay gets longer, the number of taps increases in proportion and as a result the convergence speed decreases. Hence, it is required to search for other techniques which use new filter structures that can reduce the complexity and better estimation techniques with improved convergence properties.

1.3 A Review of Hall Acoustics

The history of acoustical hall science goes back to early 1900's. It has first been introduced to the literature by W. C. Sabine [3,4] who developed an equation for calculating the parameters of sound reflections at each frequency in an architectural space. The acoustical measurements were analyzed with the aid of a graphic level recorder until the last decade. A sound source was used to excite the acoustic hall to a steady state, then the sound source was cut off and the logarithm of the decaying rms sound was plotted against time. This was the method to compute the time it takes for the sound to decay 60 dBs. This method was having some difficulties, like the beginning of the decay curve varies considerably from one excitation to another. M. R. Schroeder [5] developed a technique that eliminates this difficulty in 1965. Later Kawakami et al. [6] extended this technique. Following these principal procedures a considerable amount of research is reported in the literature [2,7,8,9,10,11,12,13].

Before discussing the acoustics of halls, let us briefly review how sound behaves in an acoustical environment. A sound wave is like a ray of light. It can be intensified, weakened, enlarged, shrunk, refracted, and more. On general terms, the acoustical sound is defined by some quantities; these are the sound pressure, the sound power, and the sound intensity. Each of these quantities has an appropriate decibel level that expresses the ratio of these quantities in relation to a reference level. The sound pressure is the pressure exerted by sound waves on any surface area measured in newtons per square meter. It is proportional to the square root of the sound energy density that is the sound per unit volume. The sound pressure level in decibels is given as

$$L_p = 20log(P_1/P_0) \tag{1.1}$$

where

L_p = Sound pressure level, in dB

P_1 = measured sound pressure, in μPa

P_0 = zero reference pressure, 20 μPa.

The sound power is the total sound energy radiated by a sound source, per unit of time. It is also called the acoustic analog of electric power. The sound intensity is the sound energy transmitted per second in the specified direction through unit area normal to this direction at the point. It is measured in watts per square meter. The relationship between sound power and intensity is given by the following equation

$$I = \frac{P}{S} \tag{1.2}$$

where

I = sound intensity, in W/m^2

P = sound power, in W

S = surface area, in square meters.

In a free field, the sound energy level decreases by 6 dB for each doubling of the distance. Inside the rooms, however, in which sound reflects from many surfaces, this level is greater than the free field level. When a vibration occurs in a room, the person present hears:

1. the original or direct sound
2. the early reflections of that sound called *echo*, and
3. the later reflections of that sound referred to as *reverberation*.

All of these variables are dependent upon the size of the room and the position of the listener.

The sound field which exists in any listening environment is the complex mixture of direct and reflected information. This field contains the direct signal path and many reflected signal paths that can exist between the sound source and the listener, as shown in Figure 1.1. These reflected signal paths can contain a single reflection or a number of reflections. Each reflected signal that eventually reaches the listener is simply a delayed and attenuated version of the direct signal. Both the delay and the attenuation are functions of the additional distance along each reflected path.

The direct acoustical sound is comprised of the sound wave that travels along the straight-line path between the source and the listener, known to be the shortest distance between the source and the listener. In the first attempts of echo cancellation, the scientists used to obtain the acoustical measurements at the listener's location as soon as the direct sound arrives. However, the acoustical echo is the reflection of that sound, which arrives the listener's location some time after the direct sound is received. One or more discrete echo signals may be detected depending on the nature of the acoustical environment. The acoustical echo signal which is to be detected as a distinct wave alone is reflected with sufficient magnitude and delay [14].

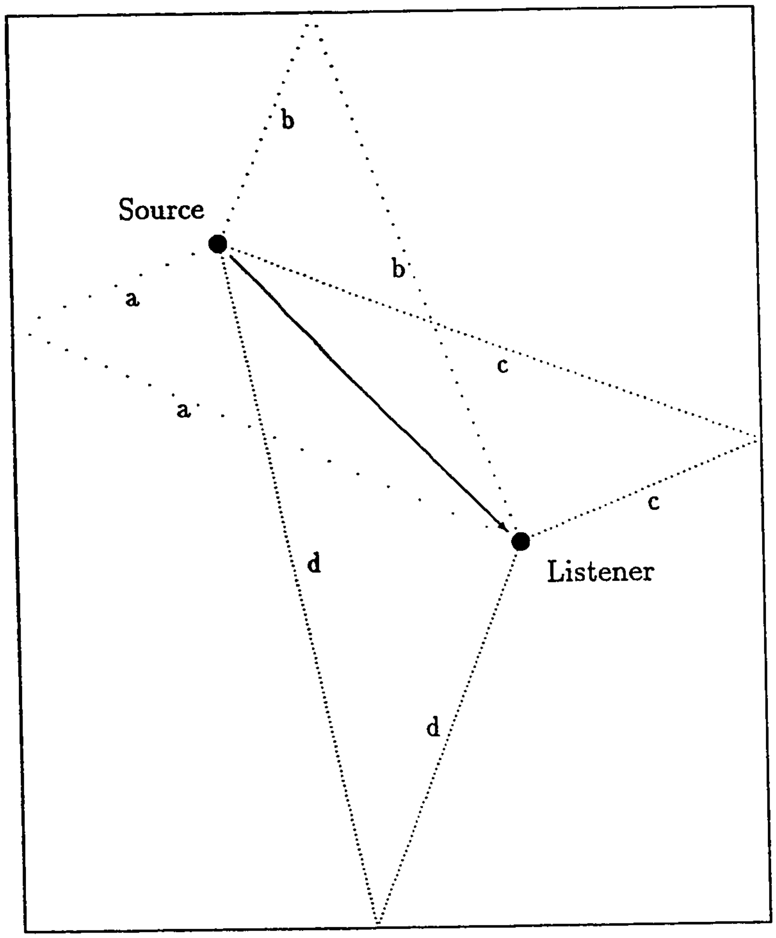


Figure 1.1: The direct and the reflected signal paths between a source and a listener.

In an ideal room, the series of reflections in the existing sound field can be described as the combination of the early reflections, the echo cluster, and the reverberation region, as shown in Figure 1.2. Every reflection is considered to be an echo, but usually the earliest reflections are defined as the discrete echoes which, in general, reduce the perception of the message in the room. The amplitude of each reflection is a function of the distance travelled to reach the listener. However, in practice, each reflection will be further attenuated since the absorption of a certain amount of energy. The decay in the reverberation region is considered as the gradual attenuation of the sound field over time, because the reflections arrive at progressively lower sound levels. The amount of attenuation varies with the frequency. Every sound field in an ideal room consists of series of reflections such as those seen in the figure, but the order and the magnitude of these reflection patterns vary from room to room, and from one location to another within the same room [14].

In an acoustical hall, however, the reflections in the sound field are comprised of the early reflections and the reverberant field. The time dependency of the whole sound field in an acoustical hall is shown in Figure 1.3. Although every echo arrives after the direct sound, the delay in such a sound field is usually taken as the time it takes for each discrete echo in the early reflection series to arrive at the listener's location, or the time at which reverberation begins. The closely spaced delays within the reverberant field are generally not considered separately [14]. In general, almost all sound system measurements are taken from the far reverberant sound field, that is the near field of an acoustic source is avoided. Most sensing

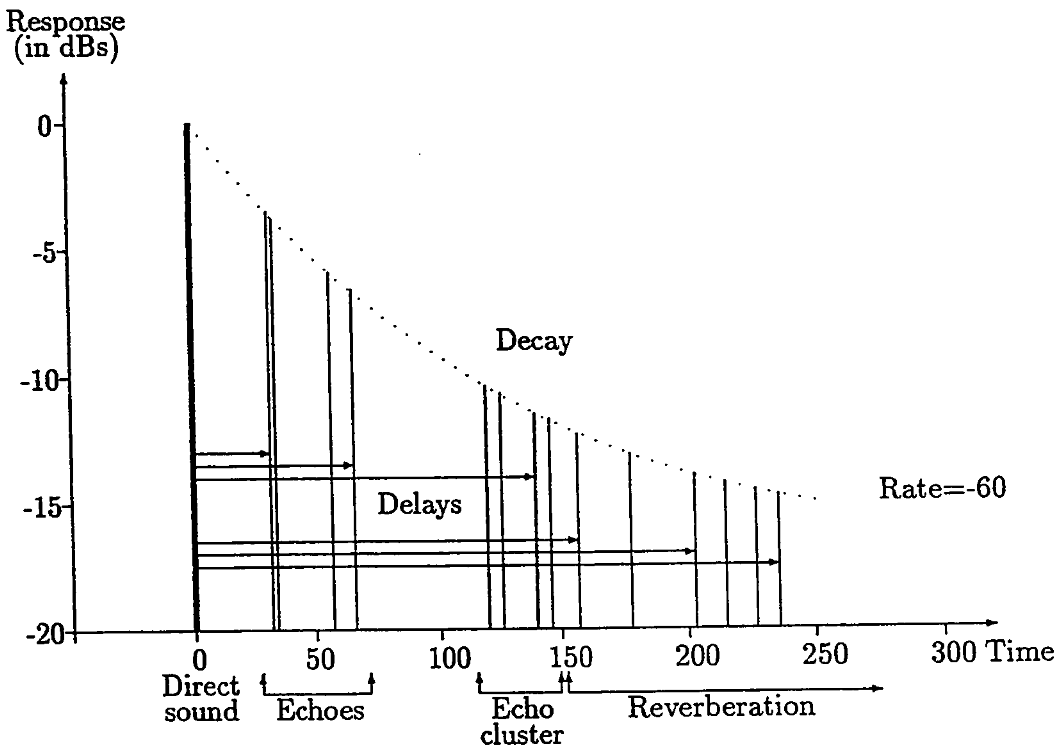


Figure 1.2: The series of reflections in a typical sound field.

$T_0 - T_{REF}$ = Signal Travel Time to Observer
$T_1 - T_0$ = Initial Time Delay (ITD) Gap
$T_N - T_{REF}$ = Natural Room Delay
$T_2 - T_1$ = 3-D Measurement Limits (T_{REF} to T_2)
L_{AMB} = Ambient Noise Level

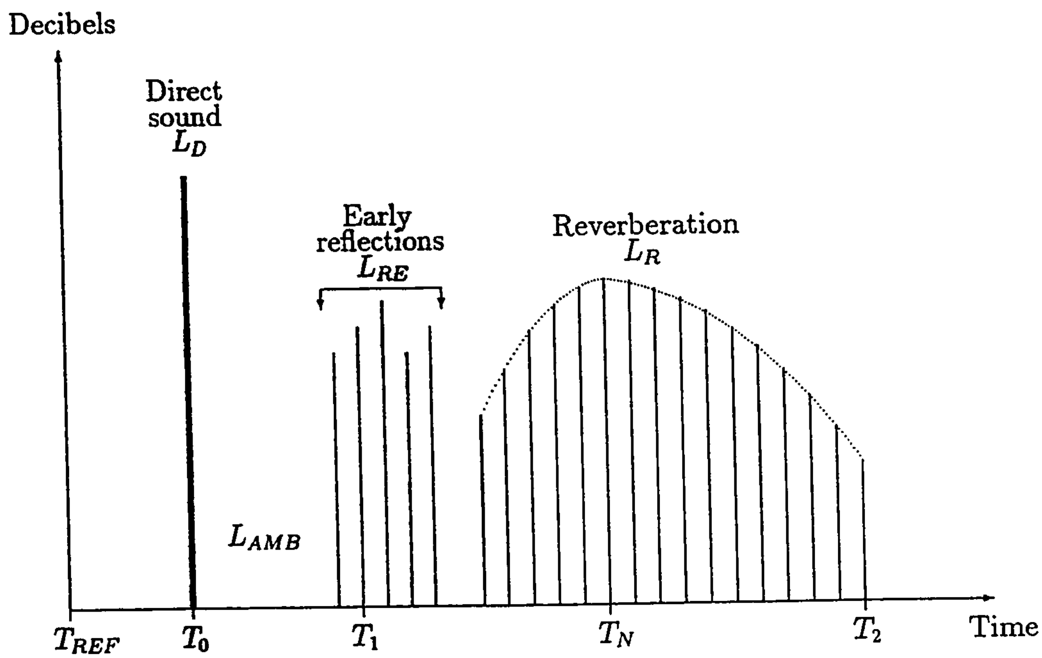


Figure 1.3: Large hall acoustic response.

microphones used with equalization measurements are placed in the far field.

1.3.1 Acoustical System of a Hall

Equalizing acoustical sound systems with the eventual purpose of increased acoustic gain and enhanced acoustic quality became an important issue in the acoustical science since late 1960's. The orientation of active research changed dramatically since the researchers were trying to see, in real time, the results of the equalizers in addition to hearing results. The principal aim became the improvement of sound quality in the sound reproduction systems as well as increasing the acoustic gain in the case of reinforcement systems.

The recommended test setup for the acoustical hall system considered in this study, Figure 1.4, basically consists of a microphone unit, preamplifier/mixer, power amplifier, and a loudspeaker system. Mixer receives signals from one or more microphones and feeds the power amplifier which transfers the signal to audience via a distributed loudspeaker system. The speech signal eventually reaches the audience via multistage from the distributed loudspeakers and from the multi reflections from the surrounding walls. Analysis of such an acoustical system is usually performed with the use of its frequency response function.

The frequency response of a typical acoustical hall system is a function of the physical characteristics of the hall in which the reflected signals are measured and additional high frequency attenuation over long distances. In a large acoustical hall, a close source will usually sound bit brighter than the same source heard over

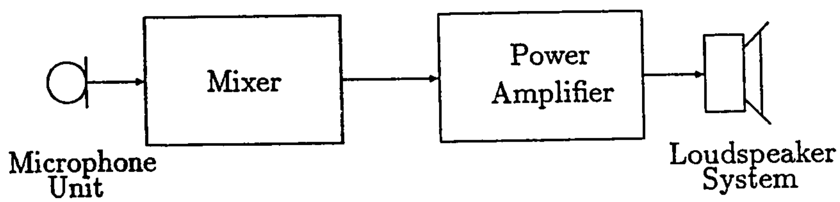


Figure 1.4: Elements of the acoustical system under study.

a greater distance, since the high frequency response decreases slightly with the distance. In addition, the reflected sound reaching the listener is usually weak in high frequencies. Therefore, the measurements taken at a distant sensing microphone lacks the high frequency content of a near field source [4,14].

The above described acoustical system is simulated through a software program to run the experiments off-line and perform the estimation and identification processes. In the proceeding chapters, an acoustical hall model, used to simulate a large acoustic hall, is described together with the transfer function derived for the hall model.

1.3.2 Transfer Function of a Large Acoustic Hall

A typical time amplitude response of a large acoustic hall was given in Figure 1.3. The response shown there consists basically of three components; the direct sound delayed by T_0, the early reflections delayed by ITD (Initial Time Delay), and the reverberant sound. In small rooms, only the direct sound level, L_D, and the early reflections, L_{RE} are present. However, in large rooms, like acoustical halls, the reverberation, L_R, is as well present together with the others. The first signal to arrive is the direct sound which is followed by the ITD gap. The end of ITD is obtained by the arrival of the first significant reflection, which is the first reflection that is higher in level than 30 dBs but lesser than L_D. The early reflections, which starts right after the ITD, are characterized by relatively distinct peaks representing the indirect or reflected focused sounds. The reverberation is basically diffused sound components causing uniformly distributed sound pressure. In small rooms, this reverberation sound field normally never appears above the ambient noise level, L_{AMB}, in the room [4].

The main factors affecting intelligibility of the speech in acoustical environments of large halls are the reverberation time, Signal to Noise Ratio (SNR), focused reflections, and alignment of the loudspeakers. The reverberation time RT_{60} is defined as the time it takes for the acoustic energy to decay 60 dBs. For a large acoustic hall it is given by Sabine equation

$$RT_{60} = \frac{0.161V}{S\bar{\alpha}} \tag{1.3}$$

where V is the volume in cubic meters, S is the surface area, and $\bar{\alpha}$ is the average absorption coefficient of the surface material. The reverberation time of a large acoustic hall is usually between 1.2 - 2.5 seconds (1.6 seconds typically) [3,4].

The effect of reverberation on the intelligibility of speech is far less than the single late high-level reflections and inadequate SNR. In a truly diffused acoustics, a high RT_{60} is not harmful to live speech until reverberation times, which are around 5 - 6 seconds at 2 kHz, are reached. On the other hand, increasing the SNR can be affected by using an additional acoustic power manually or automatically with use of an adaptive technique called Ambient Noise Controlled Amplifier [4,14].

The early reflections and multiple loudspeakers misalignment by far are the major cause of severe loss of intelligibility in the public address systems. Figure 1.5 shows the Doak & Bolt 10 percent annoyance chart which describes the trade-off between relative levels at the listener's location and the time delay. The figure

SIGNAL-DELAY CORRECTION REQUIRED IF DELAYED SIGNAL LEVEL AND TIME DIFFERENCE FALL ABOVE LINE.

SIGNAL-DELAY CORRECTION NOT REQUIRED IF DELAYED SIGNAL LEVEL AND TIME DIFFERENCE FALL BELOW LINE.

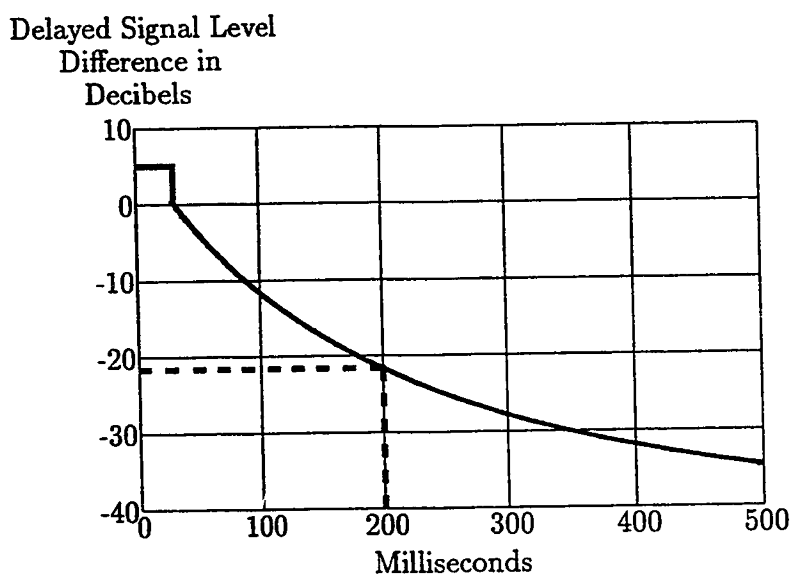


Figure 1.5: Doak & Bolt 10% annoyance chart.

shows that the effect of reflections on the perception increase with both the delay and the magnitude of the reflections. The main objective in acoustic design is to ensure that the early reflections are all under the D&B curve [4].

Traditional techniques involve distribution of drapers and reallocation of loudspeakers. In more recent techniques, the loudspeakers are fed from a number of amplifiers where the signals are fed with different delays and the gain are adjusted to ensure the D&B criteria is satisfied. The technique requires multiple amplifiers and expensive precision delay units. Adjusting the delay and power of each loudspeaker requires a lengthy procedure of trial and error.

In the simulation of the natural reverberation using the digital computer, it is shown in the literature that a delay line with feedback behaved like a reverberation process [15]. The simplest model of reverberation described by a delay line with feedback is shown in Figure 1.6. A system as shown in the figure has an impulse response like a series of echoes of decreasing amplitude separated by the delay time, and a frequency response like a comb filter, with comb spaces equal to the reciprocal of the time delay. The reverberation time of such model is given by

$$RT_{60} = \frac{3\tau}{-log_{10}(g)} \tag{1.4}$$

where τ is the delay line length and g is the feedback coefficient. In order to simulate the reverberation, one cannot use this simple structure since the high echo density (the number of echoes per second) is accompanied by a low eigentone density. To increase the echo density, however, a series of such simple networks

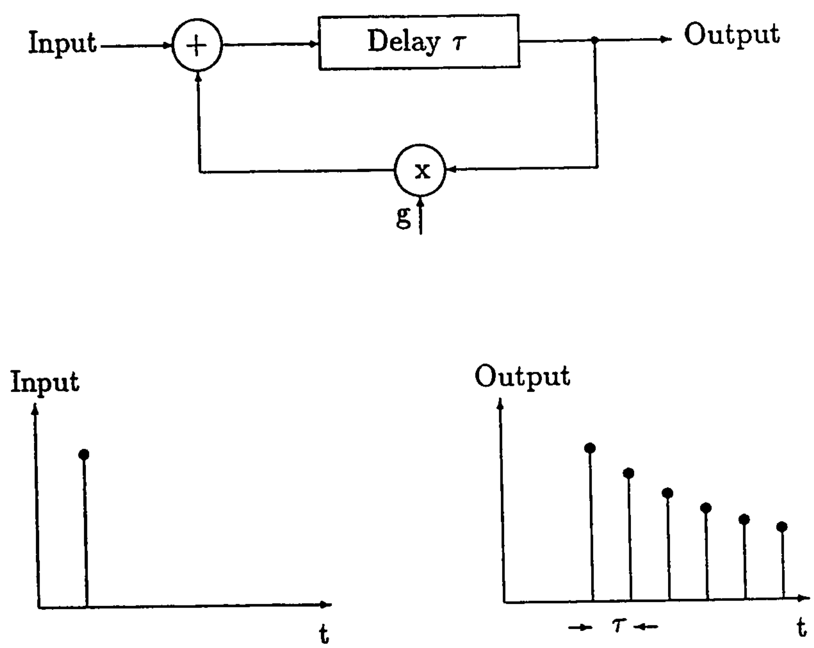


Figure 1.6: A delay line with feedback to simulate reverberation.

can be cascaded. This way the echo density will increase with the expression

$$d = t^{n-1} \tag{1.5}$$

where n is the number of sections cascaded, and d is the total eigentone density which is the reciprocal of the total delay in the system [15].

In principle, the above described system is a good approximation to real reverberation. However, this structure will let some of the eigentones from different sections to match one another. This is because of the number of discrete sections in the structure, and it will produce very strong resonances at some frequencies but it will repeatedly attenuate the other frequencies. Schroeder [5] found that each individual section could be made into an all-pass network by adding some of the unreverberated signal to the output of the feedback delay line. The main advantage of this all-pass configuration is that the frequencies that are not strongly reverberated will be passed unattenuated to the next all-pass section, where they might be reverberated. The echo density in such all-pass structured system increases as t^4, which is faster than the natural reverberation. Figure 1.7 shows an acoustic reverberation model which is modified by Schroeder and used for electronic simulation of acoustic halls. The long-term reverberation is provided by the four main loops, while the desired short-term echo density is provided by the following two all-pass networks. This way the main reverberation process does not contain a single all-pass network any longer, since all four main loops are excited in parallel. These reverberation structures assures relatively short delay times for high echo density [15].

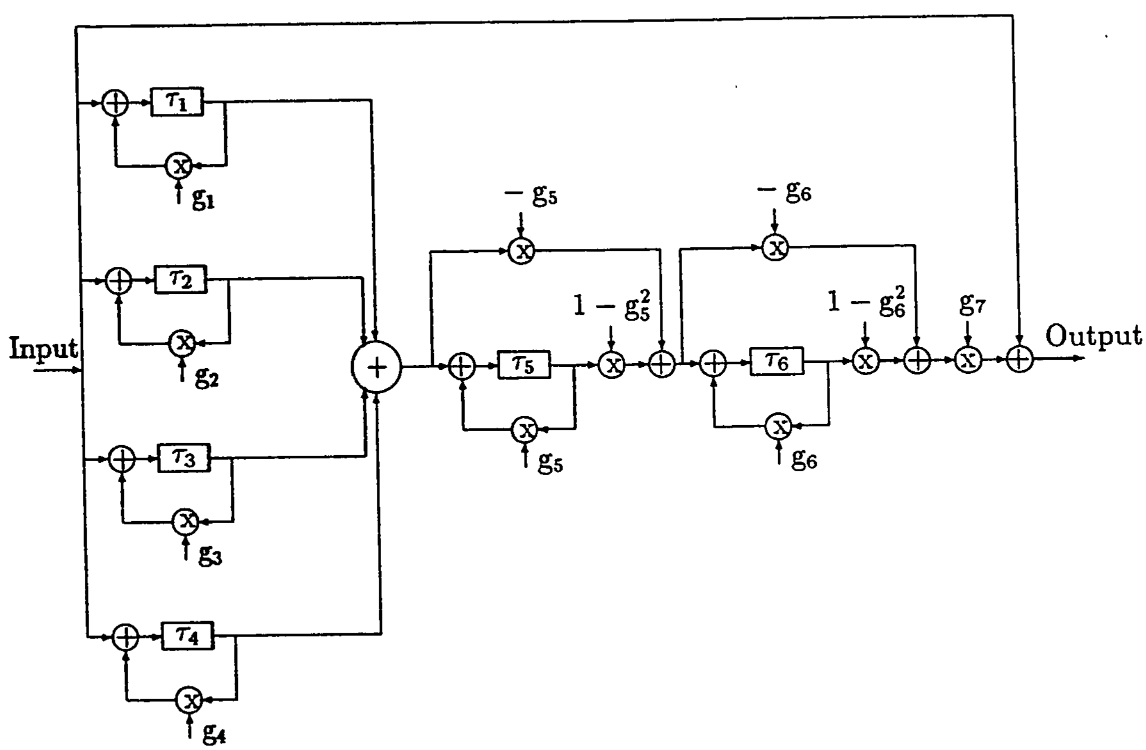


Figure 1.7: Complete reverberation model.

In summary, the transfer function of a large acoustic hall can be modeled by

$$H(z) = z^{-m_0}\frac{B(z)}{A(z)} \tag{1.6}$$

where

$$A(z) = 1 + a_1 z^{-m_{a_1}} + \dots + a_i z^{-m_{a_i}} \tag{1.7}$$

$$B(z) = b_0 + b_1 z^{-m_{b_1}} + \dots + b_j z^{-m_{b_j}} \tag{1.8}$$

where m_0 is the initial time delay, m_{a_i} and m_{b_j} are the ith and jth feedback coefficients of the corresponding block, and a_i and b_i are the delay times. Alternatively, the transfer function can also be represented by its pulse response as allows

$$H(z) = z^{-m_0}(h(z) + z^{-m_{i+1}}G(z)) \tag{1.9}$$

where

$$h(z) = 1 + h_1 z^{-m_1} + \dots + h_i z^{-m_i} \tag{1.10}$$

Here, h(z) represents the direct sound plus the early reflections, while G(z) accounts for the reverberation part. In this study, it is assumed that the perception problem is mainly due to the high amplitude peaks of the early reflections, and the part G(z) plays a secondary role. Taking this into account an adaptive design method for the acoustical echo canceller will be explored to equalize the early reflections part and obtain a desirable acoustic response.

1.3.3 Acoustical Quality of a Hall

In the real world of acoustics, especially in the large acoustic halls, there is not an ideal listening environment as in the case of an ideal room. The perceptions of a listener in an acoustical hall are influenced by the resultant sound field, specifically by the discrete echoes in this field. The acoustical designs of the halls differs from one to another, since a hall intended primarily for speech should have a shorter reverberation time than one intended for music, for example. The reverberant level and the reverberation time can both be decreased by increasing the absorption. Therefore, the optimum reverberation time is a compromise between clarity, which requires a short reverberation time, sound intensity, which requires a high reverberant level, and likeness, which requires a long reverberation time. The size of the acoustic hall and the use for which it is designed for are the two main factors in determining the optimum reverberation time [4].

In order to determine a criteria for "*good*" and "*bad*" acoustics, there has been several attempts in the literature on various acoustic halls. Beranek [3] found some subjective attributes of musical-acoustic quality that can be related to concert hall acoustics. Some of these attributes are like likeness, loudness of the direct sound, reverberant sound level, definition or clarity, etc. In designing the acoustic halls, things to be avoided can be categorized as the echoes, flutter echoes, sound focusing, sound shadows, and the background noise. A combined research by Schroeder et al. [16] concluded the following:

1. The greater the early decay time, the greater the preference for the hall, up to a reverberation time of two seconds. Above two seconds, the preference for the hall decreased with increasing reverberation time.

2. Narrow halls were generally preferred to wide ones.

3. Considerable preference was shown for halls having a high "*binaural dissimilarity*"; that is, listeners preferred dissimilarity of sound at their two ears, such as might result from a high degree of sound diffusion.

4. Halls with less "*definition*" were preferred. Definition represents the ratio of energy in the first 50 msec. to the total energy.

Cancelling out or reasonable minimizing the acoustic influence of reproduced sound waves in an acoustical sound field is, therefore, taken as the main scope of this research study.

1.4 Proposed Equalizer System and Its Motivation

The desirable configuration of the equalizer system considered in this study consists of a sensing microphone and an equalizer system between the mixer and the power amplifier as shown in Figure 1.8, where a forward path is reserved and the equalization signal can be added to it. The sensing microphone is used to receive the information on the incoming echo signal. The equalizer system which eventually be an adaptive filtering system is used to provide the required actions for the echo equalization. Such a configuration enables one to add, control, or delete the equalization signal independent of the original signal by even manual control

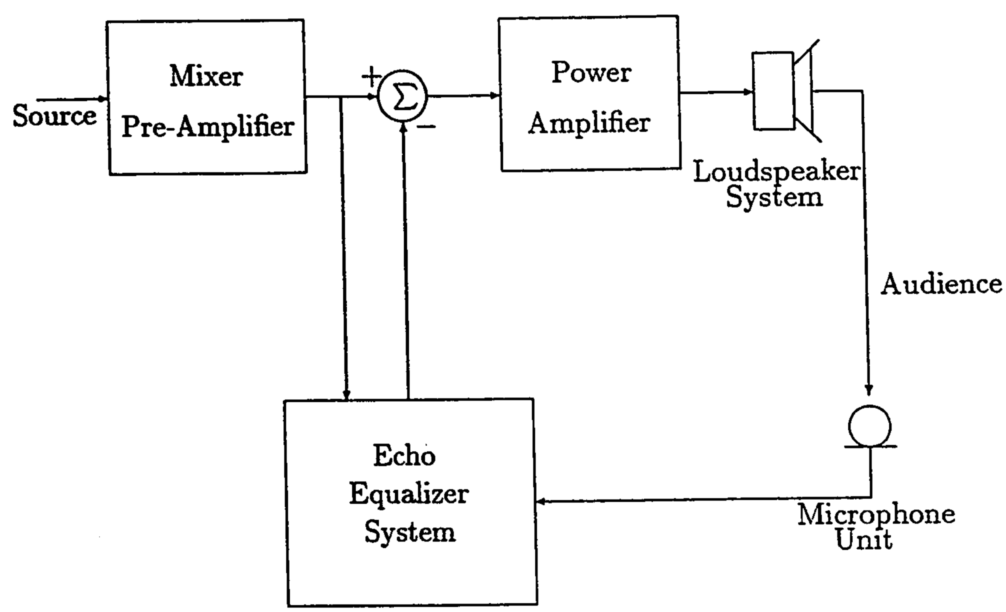


Figure 1.8: Representation of an equalizer system.

if necessary. In addition, this configuration enables the equalization process to be performed at a reduced bandwidth say 4 kHz rather than having to process the full bandwidth of speech at a sampling rate as high as 50 kHz or more.

A block diagram of the proposed echo cancellation system is shown in Figure 1.9. The signal from the mixer, typically about 0.7 RMS volt, is sampled at 8 kHz to produce the reference signal $x(k)$ which is the digitized input speech signal. At the same time the signal from the sensing microphone is also low-pass filtered by a 4 kHz antialiasing filter and down sampled at 8 kHz to produce the measured signal $y(k)$ which is the digitized output echoed signal. The input and output signals are sampled synchronously when recorded. However, the proposed cancellation system can also perform properly at reduced sampling rates. In the following chapters, internal characteristics of the proposed system will be described in details.

1.5 Thesis Outline

In Chapter 2, the acoustical echo cancellation systems; history, concepts, detailed literature review, and characteristics are discussed. Conventional techniques of acoustical echo cancellation are explained. The steps in the acoustical echo cancellation are described in details. The adaptation technique, estimation and identification algorithms used in the echo cancellation are reviewed. At the end of the chapter, a complete acoustical echo canceller system is given.

The proposed system design is presented and shown in Chapter 3. The theory and model of the proposed echo cancellation system and its equalization algorithm

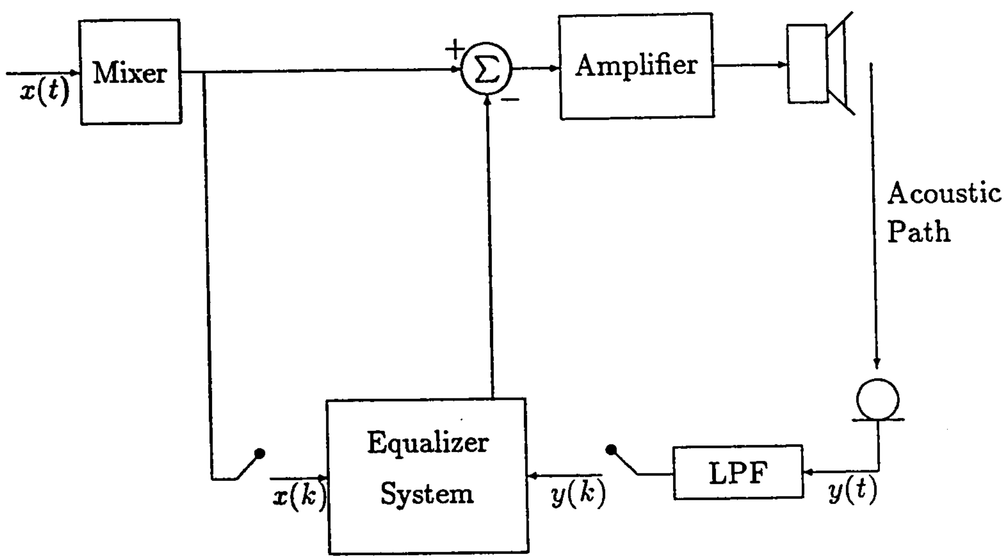


Figure 1.9: Proposed echo cancellation system.

are discussed. Mathematical model of the proposed system with all derivations is given. Advantages and disadvantages of the proposed system and also the comparisons with the available techniques are given. Finally, the difficulties encountered in acoustic echo cancellation are presented with the emphasis on the quality of perception which is the scope of this research study.

In Chapter 4, all results of the research made on the acoustical echo cancellation problem are presented and analyzed. Effects of the different estimation and identification techniques on the performance of the echo canceller are shown with the simulation examples. Performance of the proposed system is also evaluated, and the technique is validated graphically. This chapter is concluded with an overview of the problems associated with the system under study.

Finally, conclusions and suggestions for further work are given in Chapter 5.

Chapter 2

ACOUSTICAL ECHO CANCELLATION

2.1 Introduction

The acoustical echo cancellation process is a challenging and computationally complex problem. The problem of cancelling acoustical echoes in the public address systems has been studied for the last decade. In the earliest studies, the researchers approach the problem in a similar way as in the public telephone networks. However, the echo cancellation problem in acoustical hall systems is different from that in the public telephone networks mainly because of the different nature of echo paths and the duration of echoes. In the public telephone networks, the mismatch of hybrids is required to be modelled, but in the large acoustic halls, a loudspeaker-room-microphone system is needed to be modelled.

In large acoustic halls, one cannot directly judge the quality of sound without being influenced by the acoustical quality of the listening hall. In acoustic science, the conventional techniques used enable the user to obtain the desired sound mix. However, these techniques are based on improving the acoustic properties of the halls and redistribution of the acoustic power, which usually require substantial remodelling and special wall covering design at enormous cost [4].

There has been several techniques reported recently to tackle the echo problem in the large acoustic halls. Most of these techniques are not yet able to keep the acoustic influence of the hall on the reproduced sound to a reasonable minimum. This is because of the problems like lack of efficiency and/or complexity of the available techniques. Hence, the area of research to investigate efficient methods for achieving echo equalization in the large acoustic halls is still open.

Adaptive filtering techniques, as it has been extensively used in the context of many other signal processing fields, are commonly used for the acoustical echo cancellation problems. Later on, the use of such filters for the acoustic signal processing of large halls will be discussed. The estimation and identification processes which are considered to be the most important and the complex parts of the acoustical echo cancellation process will also be described in details.

2.2 Overview of the Acoustical Echo Cancellation Systems

In acoustical echo cancellation problems, as indicated before,it is required to model a loudspeaker-room-microphone system instead of modelling the mismatch of the hybrid. A simple block diagram of acoustical echo cancellation process is shown in Figure 2.1. The process basically consists of a loudspeaker system, the acoustic hall, sensing microphone, and the echo cancellation system.

The input source signal, $x(t)$, is transmitted to the acoustic hall via the loudspeaker system. The received echoed signal, $y(t)$, is then recorded through a sensing microphone and transferred to the echo cancellation system. General structure of an echo cancellation system consists of the adaptation and the cancellation processes, as shown in the figure. In the adaptation process, physical characteristics of the acoustic hall are estimated adaptively using the input and output signals. In the echo cancellation process, however, the input source signal is manipulated with the estimated characteristics of the acoustic hall to produce the signal called the echo replica, $\hat{y}(t)$. This echo replica is then subtracted from the input source signal to complete the echo cancellation process.

In such applications, the measured impulse response of an acoustical hall typically has a few thousands taps. In order to make the adaptive algorithm more computational efficient, it will be proposed to limit the number of taps to a sufficient amount. This will be clarified in the succeeding chapters.

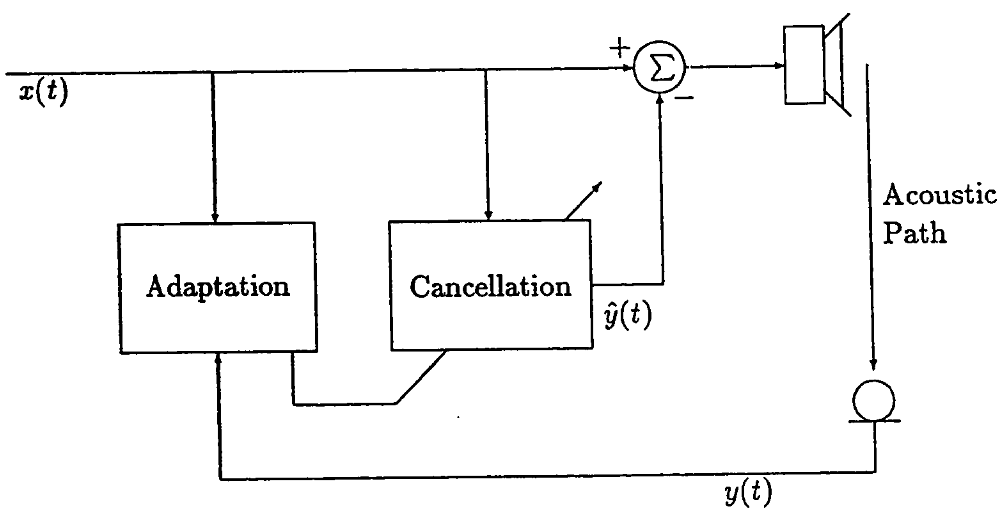


Figure 2.1: Block diagram of acoustical echo cancellation process.

In general, an acoustical echo path is modelled by a special adaptive filter structure. When the aim is adapting the system to the external environment (in the case of public address systems the input speech signal) this kind of filter implicitly estimates some characteristics of that environment. The estimates in this case are the internal parameters of the adaptive filter. The adaptation process, in general, utilizes the input and the error signals to adjust the input-output characteristic of the echo cancellation process. In this research study, the acoustical halls assumed are modelled by Infinite Impulse Response (IIR) filters and FIR filters, on the other hand the proposed echo equalizer structure will be an adaptive FIR filter structure. An IIR equalizer filter implementation is not realizable in this case, because of the stability problem of the IIR structure.

The echo cancellation process in the telephone networks deals with the cancellation of the line echoes, that is the talker echo and the listener echo [1]. In the acoustical halls, however, this process is far more different and complex from the analysis and implementation point of views. The echo cancellation problem in this case is concerned with the multiple echoes which are due to the multiple reflections from the walls and objects of acoustic hall.

The complexity and difficulty of the acoustical echo cancellation problem become obvious if the typical impulse responses of two echo paths (the acoustical echo path and the line echo path) are compared. The duration of the impulse response of the acoustic echo path is usually several times longer than that of the line echo path, and has a usual length of 100 to 400 ms. The acoustic echo path can change very rapidly at any time, for instance, due to opening door or a moving person.

In the acoustical echo cancellation process, parametric representation of the acoustic hall is described by an adaptive filter structure as discussed before. This filter, which is also referred as the acoustical echo canceller, estimates the echo signal on the basis of signal measured from the sensing microphone. The acoustical echo signal, in reality, is created by this measured signal via an unknown filter called as the echo path. In this way, the measured signal will be the combination of the echo signal and the reverberation signal carrying some of the information from the other locations of acoustic hall.

In general, an adaptive FIR filter, which can be implemented by many filter structures, is used in the parametric representation of acoustic halls. There are two filter structures which are almost universally utilized in adaptive filters. These two structures are the transversal and lattice structures. The adaptive transversal filter structure is rather direct realization of the transfer function given by equation

$$H(z) = \sum_{m=0}^{M} h_m z^{-m} \tag{2.1}$$

for a large acoustic hall in terms of delays and multipliers. This filter structure is taken as the modelling tool in this research study.

The adaptive transversal filters have been used in a variety of applications. These applications include system modelling, channel equalization, interference cancelling, etc. Sondhi et al. [1] reported that a transversal filter with over 1000 taps at 8 kHz is needed to model acoustic hall characteristics for a typical office in order to achieve a modest improvement on the reproduced sound. Some of the commercial products [2] require as many as 4000 taps at a 16 kHz sampling rate.

Implementations of the adaptive transversal filters are performed by two different methods, the time domain method and the frequency domain method. A frequency domain implementation of an adaptive filter, which has been proposed by Dentino et al. [17], requires less computation than its time domain counterpart. The circular convolution of the filter input and impulse response is employed by this frequency domain implementation. However, the conventional time domain adaptive filters employ linear convolution. The circular convolution prevents the frequency domain filter from converging to the optimum transversal filter solution attained by the conventional adaptive filter.

A different realization of the conventional adaptive filter in the frequency domain, which converges to the optimum transversal filter solution, is proposed by Ferrara [18]. It is a direct replacement for the adaptive filter, but requires substantially less computation than the time domain implementation for large filters. This method converges at the same rate as the time domain filter.

In general, the reverberation time RT_{60} of a large acoustic hall may extend to over two seconds. However, the early reflections period usually occurs within 250 ms. after the direct sound is heard. Referring to the D&B delay versus peak criteria as in Figure 1.5, it can be seen that only peaks within 30 to 250 ms. are required to be kept within the specified level. The estimation process within such a relatively small range is not an easy task because of the large number of computations involved [4]. One way of reducing this large number of computations is to use of few hundreds of taps (like 250) instead of thousands taps. In such cases, the echo reduction is possible rather than complete echo cancellation.

The acoustical echo cancellation systems generally require a large computing power due to the adaptation process. Identification of the adaptation process is a challenging problem because of :

1- the length of the impulse response and

2- faster converging algorithms are desired since changing nature of the acoustic hall.

The techniques, that use Least-Mean-Squares (LMS) type adaptation algorithms, partly meet the above requirements. Haykin [19] reported that such algorithms are known to perform badly for long impulse responses and with speech as the input signal.

The least-squares (LS) and recursive least-squares (RLS) type of algorithms increase the convergence speed. However, no one yet has shown a reasonable echo canceller having such algorithms because of the high complexity due to the long impulse responses involved. Therefore, in the literature, other methods have been explored to improve the convergence speed of LMS type algorithms. One early approach is to use of an LPC (linear predictive coding) inverse filter which aims at "whitening" the speech signal for the adaptation. This approach leads to faster convergence with little increase in complexity [7]. A simpler version of the similar idea has been proposed by Yasukawa et al. [20]. It is essentially a continuously updated first order linear predictor. Sondhi et al. [1] proposed a different approach which exploits the structure of the impulse response of the acoustic echo path and assigns different step sizes to different sections of the echo path impulse response. Thus, the impulse response samples with large values are adapted with large step size while those with small values are adapted with small step size. This way a faster overall convergence is obtained. However, the efficiency of this method depends on the priori knowledge concerning the current echo path impulse response and the ability to adjust the step size accordingly.

All of the approaches given above attempt to make compromise between the computational complexity and the convergence speed. The subband structure, which is first introduced by Furukawa [20], reduces computational complexity and at the same time provides circumstances for better convergence speed. Kellermann [7] proposed a general framework of the subband structure for the echo cancellation. The basic structure of the subband approach is given in Figure 2.2. The input and output signals of the echo path as shown in the figure are passed through identical analysis filter banks, A, producing vectors of N subband signals which are sampled at a reduced rate.

The cancellation block C, shown in the figure, forms a vector of subband signals $\hat{Y}$,

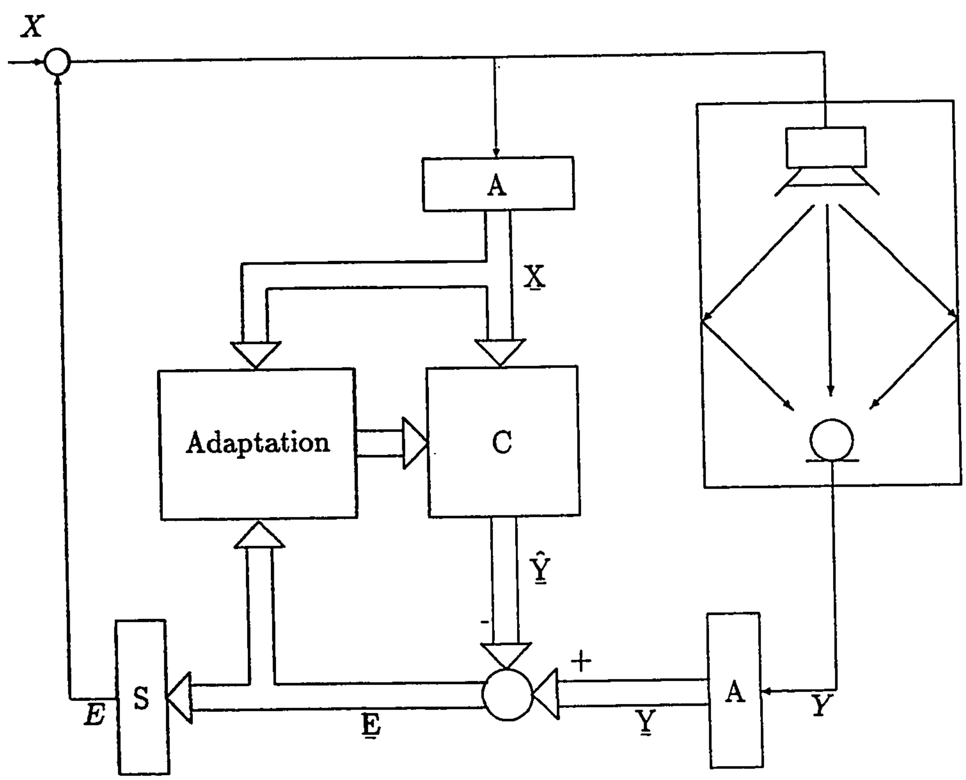


Figure 2.2: Block diagram of a subband acoustical echo cancellation process.

called the echo replica, to approximate the corresponding subband echo signals Y. The subband errors E, obtained later, are passed through a synthesis filter banks system, S, to give a full band signal E which is the error signal. This fullband error signal is then transmitted back to the remote speaker. In the adaptation block, the vector of subband error signals and input signals are utilized to adjust the input-output characteristics of the cancellation unit so as to drive the vector of error signals toward zero.

The equalization process, in the acoustical halls, basically consists of generating an echo replica, then subtracting this replica from the input speech signal. Eventually, the signal obtained is tested through the existing acoustical echo path. Construction of the echo replica is performed with the use of the filter coefficients estimated in the adaptation process. To construct the echo replica, the input speech signal is multiplied with the equalizer filter coefficients at the corresponding locations which are the locations of these estimated coefficients. This modified form of the input speech signal is referred as the echo replica.

In order to cancel out the effect of the echo, i.e. to perform the echo cancellation, the echo replica is subtracted from the input speech signal. The resulted signal after the subtraction is then processed through the acoustic hall environment for the performance evaluation of the echo equalization process. In the equalization process, possible use of windowing functions for LS techniques and weighting functions for recursive techniques may effect the speed of equalization, that is the convergence speed of the adaptation process and the real time equalization of the signal.

In the literature, more recent techniques have been reported on the acoustical echo cancellation process. Makino et al. [8] proposed an acoustic echo cancellation algorithm which is based on the variation characteristic of a room impulse response. Normalized LMS (NLMS) algorithm is used for the adaptation of the filter coefficients. The proposed algorithm, called the Exponential Step (ES) algorithm, implements a different value of step gain (feedback factor) for each tap coefficient of the canceller. The step gain values are determined proportionally to the expected variation in a room impulse response.

Nelson et al. [10] proposed adaptive inverse filters for cancelling the acoustic echoes in multichannel sound reproduction systems using the stochastic gradient algorithm with appropriate modelling delays for the adaptation. The LMS approach is used for designing these digital filters. The proposed canceller does not only equalize both the response of the loudspeakers and the listening room but also the crosstalk transmission from right loudspeaker to left ear and vice versa. Hatty [13] proposed Fast Recursive LS (FRLS) algorithms which use multirate systems like digital filter banks for cancellation of acoustical echoes. They require multiprocessor systems for parallel real time implementation. The computational complexity is reduced significantly and the instabilities of FRLS algorithms handled in a more easy fashion. It has been indicated that the convergence speed, tracking capability, and the stability are in the reasonable ranges.

A fast converging subband acoustic echo cancellation algorithm is reported by Gay et al. [11]. Algorithm is an application of Row Action Projections (RAP) which is a subset of projections onto convex sets and generalization of LMS algorithm. The proposed RAP algorithm, which can be used to improve the tracking ability of subband acoustic echo cancellers, has a useful complexity versus speed of convergence trade-off. It can use additional computational resources to speed up the convergence. This is done by allocating computational resources among the subbands in such a way as to improve the overall performance of subband echo canceller.

2.3 Steps in the Acoustical Echo Cancellation Process

In general, the acoustical echo cancellation process breaks down to the following tasks :

1- Low pass filtering and down sampling the reference signal from the preamplifier and the measured signal from the sensing microphone.

2- Estimation of the bulk delays of the digital equalizer filter $H(z)$.

3- Identification of the equalizer filter coefficients.

4- Filtering the incoming reference signal by the digital equalizer filter $H(z)$.

In order to implement the equalizer system, tasks 1 and 4 are to be performed in the real time, i.e. they have to be executed at each sampling interval by an interrupt driven routine. Tasks 2 and 3 require much more computational steps and can be computed in the background every several seconds or minutes. In the studios and theaters, in addition to equalization, task 4 could involve a high acoustic reshaping as well. In public address systems, however, achieving a reasonable perception is the first goal, i.e. reasonable echo reduction rather than echo cancellation is adequate.

In this research study, task 1 is performed through a commercial DSP board. The input speech signal from the preamplifier and the output echoed signal from the sensing microphone are first fed to the DSP board and then, as it has been indicated earlier, the digital forms of both signals are made available at 8 kHz for the later steps in the echo cancellation process. In the following sections, the major tasks of the equalization process, which are the tasks 2,3, and 4, are described in details.

2.4 Large Time Delay Estimation

The large time delay, which was first clearly identified and described by Beranek [3] as ITD given in Figure 1.3, is a fundamental acoustic hall parameter. This time delay is defined as the time between the arrival of the direct sound at a listener's location and the arrival of the first significant reflection. The term significant means the first reflection whose level is approximately close to that of the peak of the exponentially growing and decaying reverberant sound field.

Accurately estimating the large time delay in an acoustic study is one of the main step. The classical technique used in the time delay estimation is to cross-correlate the input signal with the received signal. Convolution and correlations are powerful operations in the applications of digital signal processing. The applications of these operations, however, draw heavily upon the speed of the process. Correlations are used to estimate the transfer function of a system from the spectral density relationships between input and output signals.

The term ***bulk delay*** is referred to the large acoustic time delay in this research study. The correlation process can be used to estimate the bulk delay between two similar signals in an acoustic environment. Figure 2.3 shows two signals $x(t)$ and $y(t)$ which are identical in shape but have time delay T_D between them. If these two signals are correlated the cross-correlation function $R_{xy}(t)$ would attain a maximum when $y(t)$ is delayed by an amount equal to the delay between the two waveforms. This is illustrated in the last part of Figure 2.3 where the peak of the cross-correlation function occurs at $t = T_D$, the delay between the two waveforms.

The correlation technique has several applications in digital signal processing areas, one of which is in the acoustical measurements where a measurement of time delay between transmitted and received echoed signals gives an indication of the impulse response of the existing acoustical environment. In acoustics, cross-correlating the transmitted signal and the received echoed signal is similar to the matched-filtering which results in peaks corresponding to the unknown delays. The strength of the received echoed signal is related to the amplitudes of these peaks which are weighted by the cross-correlation function.

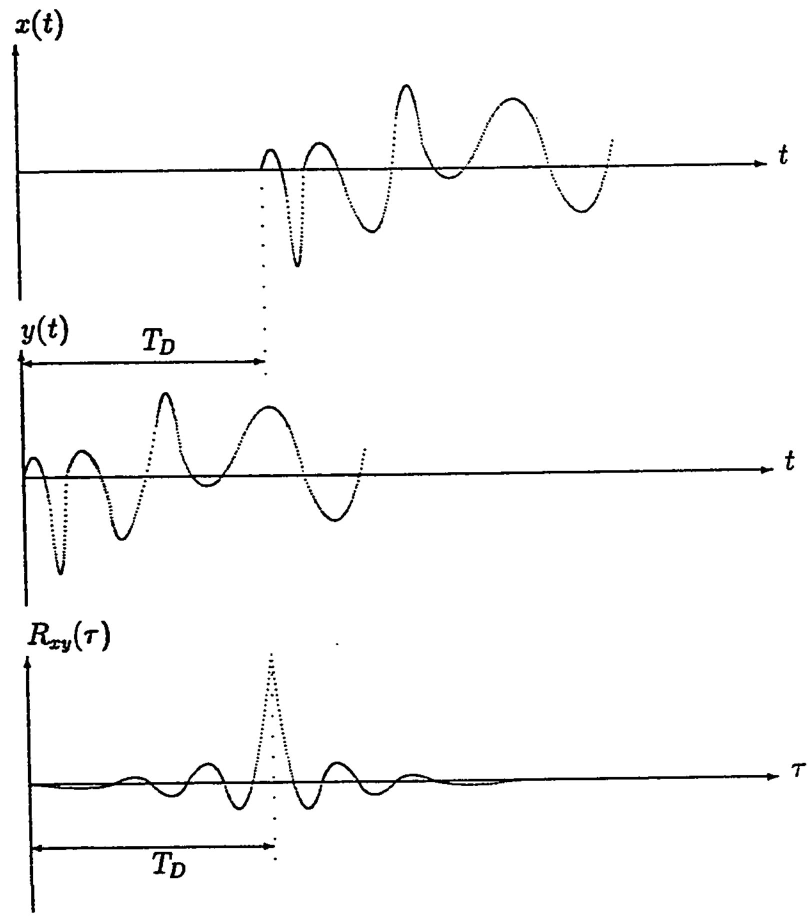


Figure 2.3: The correlation process to estimate the time delay.

In order to state the bulk delay estimation problem, let the input signal, $x(t)$, which is white, be known. The reason behind this assumption is that if the input signal is not white, for example a speech signal, then computed correlation matrix will be illconditioned. This is because of the physical properties of the speech signal. Similarities and repeatations of the speech signal, for instance, will behave like large delays in the estimation process and they will result in wrong estimates. Furthermore, let the received echoed signal, $y(t)$, be a sum of delayed versions of the transmitted signal with amplitudes representing the reflectivity of acoustic environment, distance to the walls of acoustic hall, and any number of other factors. The received echoed signal in this case can be written as

$$y(t) = \sum_{i=0}^{N} c_i x(t-i) \tag{2.2}$$

where N is the number of filter coefficients, c_i's are the amplitudes, and i represents the time delay. In general, number of sample points, N, is on order of thousands. For instance, it is required to estimate 2000 taps for an impulse response of 250 msec. with sampling rate of 8 kHz. As mentioned in the proceeding sections, the received and transmitted signals are assumed to be in sampled form with uniform spacing, so that

$$y(n) = \sum_{i=0}^{N} c_i x(n-i) \tag{2.3}$$

The continuous unknown delay and amplitude profile can now be obtained by cross-correlating the transmitted and received signal as

$$R_{xy}(m) = \sum_{k=0}^{N} y(k)x(k-m). \tag{2.4}$$

Using this correlation function the impulse response representation of the echo path of the existing acoustical environment (with peaks corresponding to bulk delays and their amplitudes) can be obtained.

Due to the restrictions in the acoustic environments of large halls, like the time range indicated above, the estimation of the bulk delays in the echoed speech signal, in general, is a complex task. It was also found that there is a large variation of the results from one segment of the measurement to the other. Fortunately, the initial bulk delay, which is usually the most significant one in such studies, can be estimated with a reasonable accuracy. Hence, any design method must be robust against the modelling uncertainty in the delay estimation [4].

In the literature, there are several approaches proposed in different signal processing fields for the estimation of time delays. Ching et al. [21] presented an adaptive algorithm for time delay estimation in the presence of a multipath reception. The proposed scheme is real time applicable since the number of adaptive parameters is reduced and the convergence time is decreased through the constraints on the coefficients which are adapted by the LMS algorithm. These constraints are restricting the filter coefficients to take on samples of a sinc function only and providing appropriate time shifts to the input signals by the adaptive filter. However, the scheme requires these constraints for the convergence. This approach has been studied for determining the location of a radiating source.

Kirsteins [22] proposed a completely different approach in which the basic idea was to look at the problem in the frequency domain. The equivalent of a time delay in the time domain is the multiplication of this delay by an exponential in the frequency domain. Thus, the frequency domain problem is one of fitting weighted complex exponentials to the spectrum of the received signal. This approach gives approximate Maximum Likelihood Estimates (MLE) with good resolution and moderate computational complexity. However, the number of different paths must be known and the spectrum of the source signal must be nonzero.

Two different high resolution time delay estimation methods, used in the underwater acoustic experiments, have been proposed by Pallas et al. [23]. These two methods are the time domain method, which is referred as the temporal method, and the frequency domain method. It has been shown that the performance of these two methods are better in comparison with the classical time delay resolution methods. Time domain method reduces the temporal spreading of the signal through matched-filtering then applies a high resolution algorithm. Good resolution and precision have been obtained in this method, however, the propagation channel must be random in order to apply this method. The frequency domain method is based on the frequency domain model of the propagation medium transfer function. This method first estimates the propagation transfer function which is the sum of undamped complex exponentials of close periods. Time delay estimation is then performed by high resolution methods like determining an L-order prediction vector describing the non-noisy signal part of the observed sequence. The frequency domain method reduces to the identification of complex sinusoids. The method looks satisfactory for delay and phase estimates.

Finally, Lee et al. [24] proposed an optimal partial-discrete algorithm for the time delay estimation. This method assumes a continuous resultant signal from which samples are obtained and processed to form a continuous profile of the delays within the known bounds. The algorithm divides the processing into three distinct stages which allow for sensitivity adjustments. The operations, in this approach, involve closed form expressions which makes the algorithm easy to implement without any approximations.

2.5 Identification of the Filter Coefficients

Accurately estimating the acoustic echo path characteristics and rapidly adapting to its variations are the basic requirements of acoustic echo cancellers. Thus, selecting an adaptive filter and an algorithm for the adaptation are two most important issues in the acoustical echo cancellation systems.

The next step after properly estimating the large time delays is to identify the coefficients of the adaptive filter to be used in identification of acoustic echo parameters. This is usually computationally expensive process since it is required to estimate a set of few thousands of tap coefficients. Identification process has also to be performed in a limited time due to the restrictions on the real time implementations. Thus, it is necessary to use an adaptation algorithm which is stable and has a fast convergence speed.

2.5.1 Adaptive Filter Structure

In digital signal processing, a common technique used for smoothing a data sequence is to take a simple weighted average of $N+1$ adjacent input values to produce each output value. A casual version of this filtering operation is thus described by the difference equation

$$y(k) = \sum_{n=0}^{N} a_n x(k-n) \tag{2.5}$$

where $x(k)$ is the input signal and $y(k)$ is the output signal. This difference equation can be implemented nonrecursively as shown in Figure 2.4 which is called as the transversal filter. The corresponding transfer function is simply

$$H(z) = \sum_{n=0}^{N} a_n z^{-n} \tag{2.6}$$

and the impulse response, $h(n)$, is obtained directly from $H(z)$ or from the block diagram as

$$h(n) = \begin{cases} a_n, & n = 0, 1, ..., N \\ 0, & \text{otherwise.} \end{cases}$$

Impulse response of this filter has nonzero values only for a finite duration, such filters are called FIR filters. Usually, FIR filters are implemented nonrecursively, but recursive implementations can also be generated.

The modelling of an acoustical echo path by an adaptive transversal filter of order N can be given as the block diagram shown in Figure 2.5. The Transversal Filter, TF, which is one of the essential signal processing structure, consists of three basic components as from Figure 2.4. These three components are the unit delay elements,T, multipliers, $a_i s$, and an adder. The number of delay elements, usually

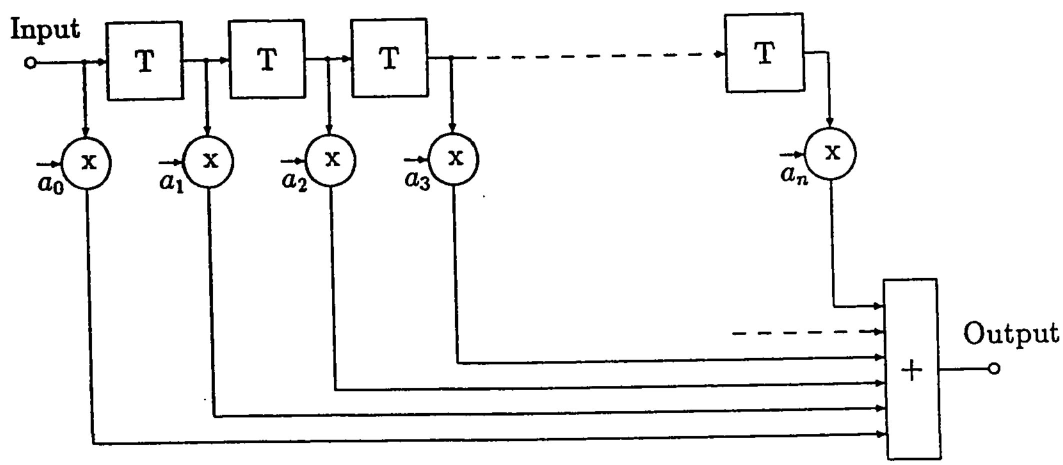


Figure 2.4: A transversal filter.

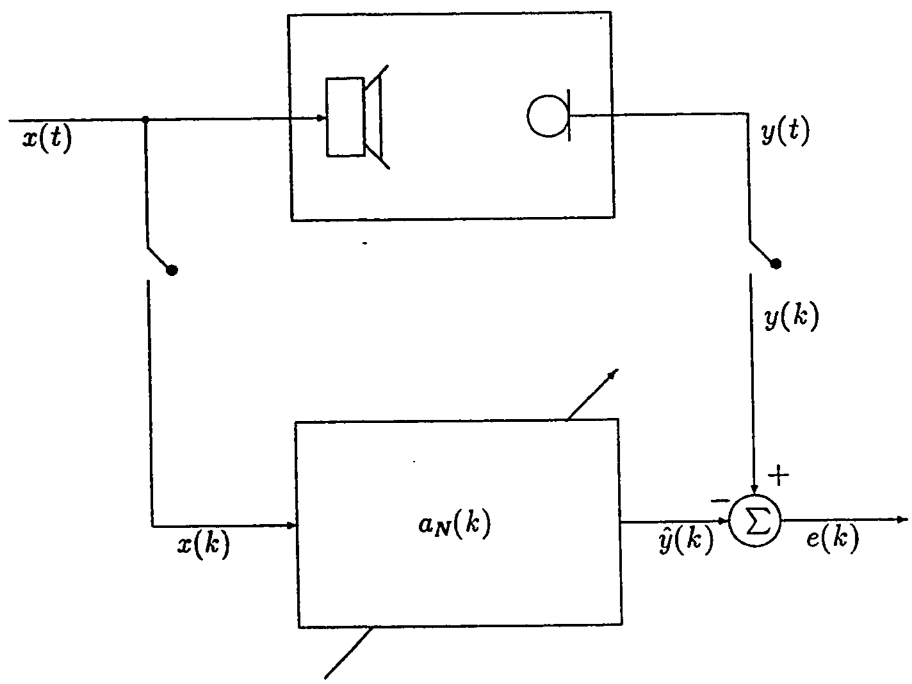


Figure 2.5: Modelling of acoustical echo path by an adaptive transversal filter.

referred as the order of the filter, used in the filter determines the duration of its impulse response. Each delay element is identified by the unit delay operator T as shown in the figure. The input signal is delayed by a chain of N stages, each stage delaying the signal by amount T. Each stage output or the tap input is multiplied by the corresponding filter coefficient a_i which is also referred as the tap weight. These coefficients must be controlled adaptively for maximum echo cancellation. The overall filter output is obtained by summing up the individual multiplier outputs.

The adaptive FIR filter (the tapped delay line), shown in Figure 2.4, is by far most widely used in acoustical echo cancellation systems. The output of such a filter is called the finite convolution sum in the sense it convolutes the finite impulse response of the filter with the filter input. The major drawback in an acoustic study is that as the echo duration becomes longer, the number of taps increases proportionally and the convergence speed decreases. In the acoustical echo cancellation systems, echoes are generated by sound waves with a long propagation delay. Therefore, the number of required taps is on order of several thousands. The difficulty of performance and complexity can partly be solved out with the FIR implementations. The optimum value of the filter coefficient vector $a_N(k)$ shown in Figure 2.5 in the least squares sense is determined by the normal equation

$$\mathbf{R}(k)a_N(k) = \mathbf{T}(k) \tag{2.7}$$

where n = 1,2,3,...,N (N is the order of filter) and

$$\mathbf{T}(k) = \sum_{i=1}^{k} \mathbf{X}(i)y(i) \tag{2.8}$$

$$\mathbf{R}(k) = \sum_{i=1}^{k} \mathbf{X}(i)\mathbf{X}^T(i) \tag{2.9}$$

$\mathbf{T}(k)$ and $\mathbf{R}(k)$ are called as the cross-correlation vector and the correlation matrix respectively and k is the discrete time index. Derivation of the above formula and the related details will be discussed in the following sections.

In the adaptive transversal filter structure, the tap-input vector $\mathbf{X}(k)$ contains the past N samples of the input speech signal $x(k)$, which is the reference signal, that is

$$\mathbf{X}(k) = [x(k), x(k-1), ..., x(k-N-1)]^T \tag{2.10}$$

The echo replica, $\hat{y}(k)$ which is basically the output of the adaptive transversal filter, is obtained by convoluting a measured impulse response of the hall $a_N(k)$ with the received input signal $x(k)$. The echo replica is then subtracted from the real sampled echo $y(k)$ to get the residual echo $e(k)$ which is the error.

In the literature, the research is currently going in several directions in an attempt to find new filter structures that reduce the complexity and improve the performance. One approach is to adapt IIR filters [1]. The reason behind using IIR is that this filter structure is more suitable if the echo path can be modeled by a combination of poles. The important points in this case are the stability and the estimation of filter coefficients that provide the optimum overall performance.

2.5.2 Adaptation Algorithm

The adaptation process in acoustic signal processing is concerned with identifying certain characteristic parameters of the observed speech signal. Based on these parameters an error measure is computed which is used to adjust the acoustic signal processing (in echo cancellation it is the input/output relationship) so as to minimize the error. The sequence of steps required to minimize the error measure is the adaptation process.

In general, there are two basic categories of adaptation algorithms used for the acoustic echo cancellers, namely the Least Squares (LS) algorithm and the Least Mean Squares (LMS) algorithm. The LS algorithm is based on the information of the past reference signals and the corresponding echoes. This algorithm determines the tap coefficients that minimize the squared error summed over time. Sometimes exponentially decreasing weights are assigned to the past error signals for gradually fading out the past data and enabling the finite dimension arithmetic.

In order to describe the LS algorithm for the adaptive FIR filter, the echo replica, $\hat{y}(k)$, can be expressed using the reference input signal, $x(k)$, as follows

$$\hat{y}(k) = \sum_{n=0}^{N-1} a_n(k)x(k-n) \tag{2.11}$$

where $a_n(k)$ is the estimate of the acoustic echo path impulse response at time k and N is the tap length. The estimated residual echo or error signal is then

$$e(k) = y(k) - \hat{y}(k) \tag{2.12}$$

where $y(k)$ is again the received echoed signal. The criterion function $D(k)$ is defined in the LS algorithm as

$$D(k) = \sum_{l=0}^{k} e^2(l). \tag{2.13}$$

The tap coefficients are adapted to minimize $D(k)$ sometimes with use of a weighting function. The following matrix equations are obtained by taking the derivative of $D(k)$ and setting to equal to zero;

$$\mathbf{T}(k+1) = \mathbf{T}(k) + \mathbf{R}^{-1}(k)\mathbf{X}(k)e(k) \tag{2.14}$$

$$\mathbf{R}(k) = \mathbf{R}(k-1) + \mathbf{X}(k)\mathbf{X}^T(k) \tag{2.15}$$

where $\mathbf{T}(k+1)$ is the correlation vector, $\mathbf{R}(k)$ is the correlation matrix, and $\mathbf{X}(k)$ is the tap input vector. The adaptation process of the LS approach involves solution of a normal equation which is determined by the correlation vector and the correlation matrix. The tap coefficients adaptation is performed in an attempt to minimize the error which is the difference between the input and output signals.

The major advantage of the LS algorithm is the fast convergence, irrespective of the correlation characteristics of the input signal. However, it requires computation of the inverse function for obtaining the optimum coefficient values and results in complex implementation.

The LMS algorithm uses a different criterion function which is the expected value of the squared error. In the case, the taps are adapted according to the stochastic steepest descent algorithm. The criterion function can be expressed as

$$D(k) = E[e^2(k)] \tag{2.16}$$

In practice, the instantaneous value of the squared error is used in place of the normal equation and the tap coefficients are controlled using the derivative of $D(k)$ with respect to the tap coefficients. The derivative of $D(k)$ is then

$$\frac{dD(k)}{da_n(k)} = -2e(k)x(k-n) \tag{2.17}$$

Thus, the recursive tap coefficient adaptation becomes

$$a_n(k+1) = a_n(k) + 2\mu e(k)x(k-n) \tag{2.18}$$

where μ is a constant. The value of μ plays an important role in determining the convergence speed, stability, and the residual error after convergence. The algorithm convergences faster for a larger value of μ, but it results in a larger value of residual error. Thus, the algorithm tends to instability. The optimum range of μ in related applications vary from 0 to 1 [2,19,25].

The LMS algorithm is widely used due to its comparatively easy implementation and its well-established stability characteristics. The major drawback of the LMS algorithm is the dependence on correlation of the reference input signal. When the input signal is a human speech, which is highly correlated, the convergence speed slows down. In the applications of acoustic echo cancellation, where a large number of taps are involved, it is necessary to use a whitened training signal or a pre-whitening filter, like lattice structure or linear predictive filter.

There are several techniques [2,8,9,11,12,13,25] used in the identification of digital adaptive filter coefficients. These coefficients have a vector form as given in the normal equation. The solution of this equation results in the coefficients of the adaptive filter or, in general, the tap coefficients. A graphical representation of these coefficients is given in Figure 2.6. This vector set of coefficients, as indicated earlier, gives the impulse response of the acoustical echo path.

2.6 Equalization Process

The first meaningful sound system equalization process is performed back in late 1950s [4]. Equalizing sound reinforcement systems with the end purpose of increased acoustic gain and enhanced acoustic quality became universal with the early introduction of real time equalizers. Since then number of techniques have been reported for the improvement on the design and analysis aspects of the equalization process. This was the start of research on the equalization process in the acoustical halls of the large halls.

The term equalization describes a signal processing process in which the primary function is to modify the response of the signal being processed. In acoustical systems, the basic purpose of this process is to compensate for any system deficiency so that the equalization process delivers a reasonable replica of the system input. It is, therefore, necessary to select an appropriate digital filter for the equalization. There are several filter design techniques used for the acoustic echo equalization. In this section, AutoRegressive (AR) filter, Moving-Average (MA) filter, AR-MA filter, and IIR filter structures will be discussed briefly.

IIR and FIR filter design methods are based on the mathematical theory of approximation. Other design methods, however, are based on the statistical modeling of

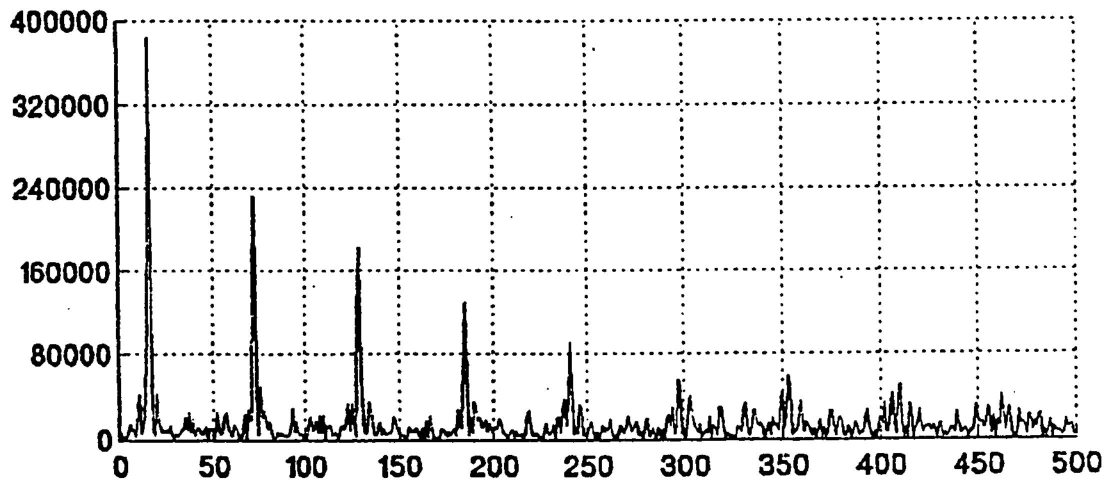


Figure 2.6: Graphical representation of the tap coefficients of an acoustic echo path.

spectra. In these methods, generally, the least-mean-squares error criterion is employed. They are especially appropriate when the filter design or model is derived from experimental data having random fluctuations.

The AR method or all-pole filter modeling is the basic approach used in all other filter structures discussed in this section except the IIR filter structure. A simple model of the AR filter structure can be given as

$$H(z) = \frac{1}{A(z)} \tag{2.19}$$

where

$$A(z) = a_0 + a_1 z^{-1} + \ldots + a_N z^{-N}. \tag{2.20}$$

In the time domain, using the notation $a(n) = a_n$, the following can be written

$$h(n) * a(n) = \delta(n) \tag{2.21}$$

with this equation, it is understood that $A(z)$, which is the FIR filter structure, is the inverse filter for $H(z)$. In this case, the filter $A(z)$ whitens $h(n)$ to produce $\delta(n)$. Using the $2N+1$ consecutive values for $h(n)$ beginning at $n = -N$, the equation 2.21 provides $N+1$ linear equations which can be written in matrix form. This matrix form is used to solve for $a(n)$ which are the filter coefficients. Covariance and auto-correlation methods are used to obtain accurate values for $a(n)$.

The MA method or all-zero filter is another name for FIR filter. The MA filter model is of the form

$$H(z) = B(z) \tag{2.22}$$

where

$$B(z) = b_0 + b_1 z^{-1} + \ldots + b_N z^{-N}. \tag{2.23}$$

An Lth order AR model can approximate any desired filter with an arbitrary accuracy if L is made sufficiently large. Thus, it is desired to have

$$B(z) \approx \frac{1}{A_L(z)} \tag{2.24}$$

or, in the time domain,

$$a_L(n) * b(n) = \delta(n) + e(n) \tag{2.25}$$

where $\frac{1}{A_L(z)}$ is a high order AR model of the measured impulse response with $L \gg N$ (i.e. the order of $A_L(z)$ is larger than that of $B(z)$), and $e(n)$ is the approximation error. The estimates of $b(n)$ can now be obtained such that $\frac{1}{B(z)}$ is an AR model of the FIR filter $A_L(z)$. To compute $b(n)$, the auto-correlation method may be directly applied to the data sequence $a_L(n)$ since $A_L(z)$ is FIR.

The general AR-MA filter structure, which is also called pole-zero model, is more complicated than the separate AR and MA filter models. The simplest form of this model is based on equation error measures having no directly meaningful interpretations. The general AR-MA model is given as

$$H(z) = \frac{B(z)}{A(z)}. \tag{2.26}$$

Linear equations may then be produced for $a(n)$ and $b(n)$ by writing equation 2.26 in the time domain as

$$h(n) * a(n) = b(n) + e(n) \tag{2.27}$$

where $e(n)$ is simply the equation error. Forcing $e(n) = 0$ for $n \leq N$, a vector form of the equation 2.27 can be obtained. The covariance method provides the estimate of $a(n)$ which is later used to obtain the estimate of $b(n)$.

A digital filter whose impulse response never decays to exactly zero is said to have an infinite impulse response and called an IIR filter. The most common implementation of such filters is to use the difference equation, a simple form of such equation with a single coefficient can be given as

$$y(n) = x(n) + ay(n-1) \tag{2.28}$$

where $x(n)$ is the input and $y(n)$ is the output signal, a is the multiplier, and 1 indicates the unity time delay. The model of this system is given as

$$H(z) = \frac{1}{1 - az^{-1}}. \tag{2.29}$$

General form of this model, that is for N number of coefficients, can be derived using the full form of equation 2.28. As opposed to FIR filter case, IIR filter is usually implemented recursively. The important properties of IIR filters are as follows

1- they are not always stable (they may have poles outside the unit circle),

2- they cannot be linear phase,

3- if the poles are closed to the unit circle, large changes in the frequency response can be caused by small changes in any of the coefficients.

2.7 Complete Acoustical Echo Canceller System

An acoustical echo canceller, as mentioned earlier, identifies the impulse response between the loudspeaker system and a microphone to produce an echo replica which is than subtracted from the real echoed speech signal. Due to the variation in the impulse response of the acoustic hall, an adaptive FIR filter must be used to identify the coefficients of the impulse response. The acoustical echo cancellation algorithm must also have the capability of performing the real time operations.

A complete acoustical echo canceller system, proposed by Sikorav [9], is shown in Figure 2.7. The system consists of an acoustical room, in which a loudspeaker and a microphone placed at two different locations, a model block for the acoustic echo path, and an adaptation block for incorporating the variations in the impulse response. Acoustical coupling in the acoustic room is also considered.

The proposed echo canceller synthesizes, in real time, a replica of the acoustic echo path. In the next stage, this replica is subtracted from the output signal of the room which is obtained by the echo signal and the local speech signal (i.e. the input signal). This way Sikorav tried to isolate the output signal as an distinct echo signal only without having any information on the direct input signal. Later, he used this echo signal in the cancellation process, treating it as an echo replica. Sikorav used a very large form of a transversal filter to model the acoustic echo path. This filter is used to simulate the impulse response of the linear system formed by the loudspeaker and microphone.

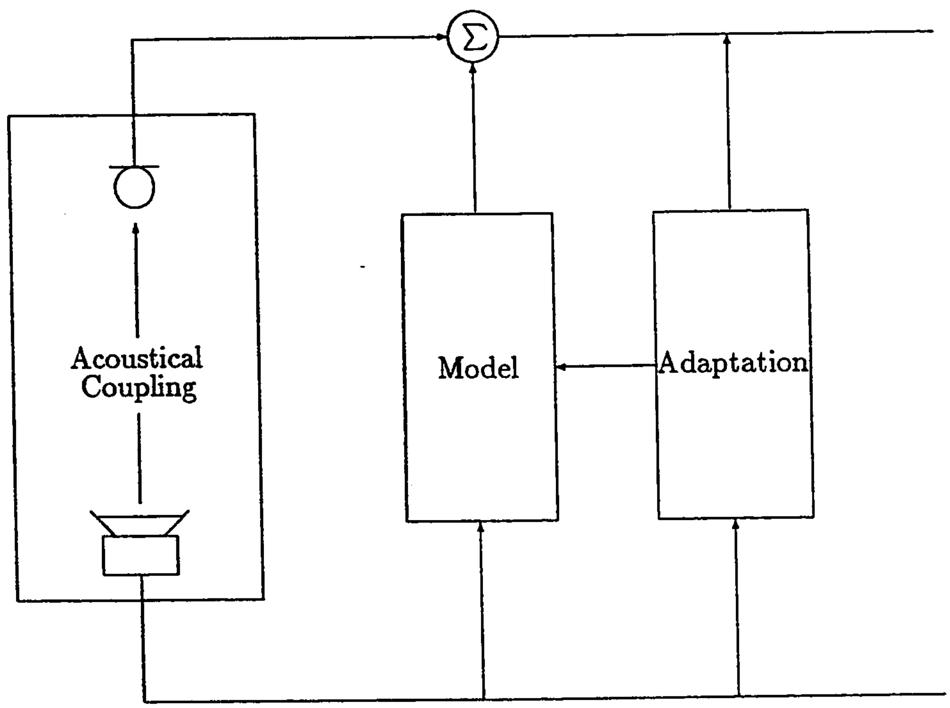


Figure 2.7: Block diagram of a complete acoustical echo canceller.

In the adaptation process, Sikorav used the LMS algorithm to improve tracking of non-stationarities in the impulse response of the acoustic room. He aimed at introducing a geometrical information which identifies the moving obstacles in the room. He concluded that the deformations of impulse response, which are due to different setups in the acoustic room, are not uniform and tracking can be improved by simple modification of the LMS algorithm that gives similar results as spectral methods like those used in subband analysis.

The echo cancellation system considered in this research study will assume that the acoustical hall is modelled by an IIR or an FIR filter and the equalization process is performed by an adaptive FIR filter. The acoustical hall models used and the adaptation algorithm will be described in details in the following chapter. However, different techniques in the adaptation process will be considered for the comparison reasons. These different techniques will be studied for the same acoustical hall model using the same input speech signals.

2.8 Summary

The acoustical echo cancellation systems have basically three components, the estimation of large time delays, identification of the adaptive filter coefficients, and the equalization process. These components are described in details together with the related literature research available on the related topics. The techniques proposed on the cancellation of acoustical echoes are explained and the advantages/disadvantages of each technique is described. The research on proper cancellation of the acoustical echoes are still open for the later improvements and for the verifications.

The acoustical echo canceller system proposed in this study will be described in the next chapter. This chapter will basically cover model of the assumed acoustic hall system, the estimation and identification theories, and the algorithm used for the echo equalization. The following chapter will demonstrate the simulation results and the analysis made on the different estimation and identification techniques.

Chapter 3

PROPOSED ECHO CANCELLATION SYSTEM

3.1 Introduction

In this chapter, the proposed acoustic echo cancellation system is presented in detail. The complete structure of the acoustic echo cancellation system used in this study is given in the following section and each part of this structure is explained in details in the corresponding subsequent sections. This system is designed to improve the perception of speech in large acoustic halls.

The system design concept is described in the first section of the chapter. In the following sections, issues for modeling the acoustic hall are explained in steps. The acoustic echo equalizer filter is then described. Finally, advantages and disadvantages of the proposed echo cancellation scheme are discussed.

3.2 System Design Concept

The design procedure for an acoustic echo cancellation system can be divided into two distinct steps. These two steps are modeling the acoustic hall and designing a proper echo equalizer for this hall. The acoustical echo cancellation process, considered in this study, has the block structure as given in Figure 2.1.

In the first step, which may also be called as the adaptation process, characteristics of the existing echo path is determined by using a training signal, $x_T(t)$, at the input. These characteristics of the acoustic hall are modeled by using an IIR digital filter as indicated before. After identifying the acoustic hall model, the filter coefficients obtained are transferred to the second step which is the cancellation process. In the cancellation process, as seen in the Figure 2.1, the original input signal, $x(t)$, is filtered through the digital equalizer filter, coefficients of which are determined in the adaptation process, to obtain the echo replica, $\hat{y}(t)$. Eventually, the echo replica is added to the input signal and both are fed to the acoustic hall via loudspeakers for equalizing the acoustic echoes.

The equalizer filter used is an FIR filter since the hall is modeled by an IIR filter. A general block structure of the proposed echo cancellation system is given in Figure 3.1. In the figure, the block with $H_h = \frac{1}{1+H}$ represents the acoustic hall model which is an IIR filter model. The block with H_c, however, represents the echo canceller which is an FIR filter model. H_h and H_c are different digital filters, however, they have the same filter coefficients.

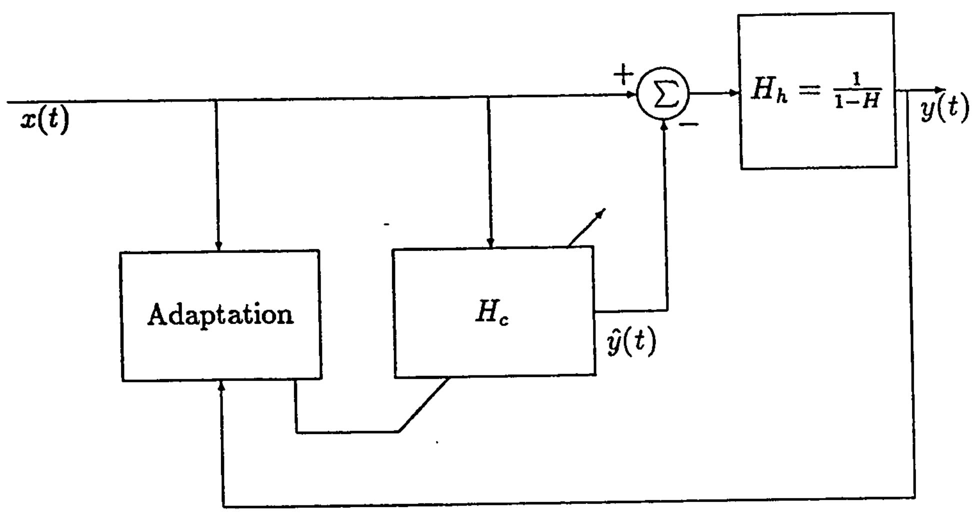


Figure 3.1: General structure of the proposed echo cancellation system.

The adaptation process is assumed to be performed only at the beginning of equalization process and at certain times in background during the equalization. It is necessary to perform the adaptation process at certain interval of times since the echo path transfer function varies due to the physical changes in the acoustic hall. During the normal operation of the echo canceller the adaptation process is performed if the error between the reference input signal and the received output signal is higher than the threshold set for the quality of perception. This error is computed using the LS error criterion. The training signal, $x_T(t)$, used at the beginning has desired characteristics since it is used for estimating the large time delays. However, during the equalization process, this signal can again be used or instead a different training signal can be used for the estimation process.

The echo cancellation system described here is capable of handling the variations in the acoustic hall environment. Thus, the equalizer filter is called an adaptive FIR filter. The characteristics of this filter and adaptation and equalization processes will be described in the following sections.

3.3 The Acoustic Hall Model

Modeling process of a large acoustic hall consists of two significant stages, namely estimation of the bulk delays and identification of the digital filter coefficients. These two processes usually require lengthy and complex computations since length of the transfer function of the acoustic echo path is on order of several hundreds of milli seconds. To model a digital filter for such transfer function, therefore, requires estimation of a few thousands of filter coefficients.

The estimation process consists of determining locations of the bulk delays which are referred as the early reflections in the acoustic science. It is necessary to determine these locations accurately since the perception of messages in the large acoustic halls is mainly reduced because of these reflections. Correlation techniques are used to estimate the bulk delays. In this study, because of its simplicity, cross-correlation technique is used to obtain estimates of the bulk delays.

Estimation of the bulk delays is followed by the process of identifying the acoustic hall digital filter coefficients. In the identification process, estimated bulk delays are used to form a vector of coefficients, the tap coefficients vector. These filter coefficients are assumed to represent model of the acoustic hall. To determine accurate values for the tap coefficients, LS method is used to obtain a vector form normal equation. This equation is in form of a set of linear equations which is solved to identify the tap coefficients.

Acoustic hall echo path model is assumed to be modeled by an IIR digital filter. The discussions on this echo path model and on the estimation and identification processes performed using this model will be described in the following sections.

3.3.1 Model of the Echo Path

The acoustic echo path of a large hall can be described by the acoustic reverberation models given in Chapter 1. An IIR model of such an echo path is described by the transfer function given in equation 1.4. This IIR model approximates both the early reflections part and the reverberation field of the acoustic environment. In this model, attenuation of the major echo signals is achieved by the decaying coefficients, a_is, the reverberation region is formed by the coefficients b_is, and m_{a_i}s - m_{b_j}s represent repeatations of the existing echo paths.

In this study, however, it is assumed that the perception problem of the speech message is mainly due to the high amplitudes of the early reflections. Since it is interested to refer to dominant early reflections, the general IIR model could be simplified to

$$H_h(z) = z^{-L_0} \frac{1}{A(z)} \tag{3.1}$$

where

$$A(z) = 1 + \sum_{i=0}^{N} h_i(z) z^{-L_i} \tag{3.2}$$

This is an IIR filter configuration which delays an input signal by amount of L_0. The coefficients h_is are the coefficients to be identified and L_is represent the delay lines.

The form of $A(z)$ can be chosen to represent either a simple or a complex acoustic model. The transfer function, $H_h(z)$, can also be represented by its pulse response as

$$H_h(z) = z^{-L_0}(1 + h(z)) \tag{3.3}$$

where

$$h(z) = h_{o_0} z^{-L_0} + h_{o_1} z^{-L_1} + ... + h_{o_i} z^{-L_i} \tag{3.4}$$

where the set of coefficients, h_{o_i}, is the original set of identified coefficients. This set of coefficients are updated in the adaptation process by the LS method to compromise the changes in the acoustic echo path. Other classical acoustic models can also be used to generate the echoed signals which are used for the estimation and identification processes. In the following chapter, the computer program code and the performance analysis of the derived acoustic hall model for the analysis are described in details.

3.3.2 Bulk Delay Estimation

The first step in modeling a digital filter for a large acoustic hall is the estimation of bulk delays. The estimation process basically consists of identifying positions of the bulk delays. In general, it is a lengthy process that requires large number of computations. Therefore, the estimation process is usually performed at the beginning of the equalization process using the $x_T(t)$ signal.

The correlation techniques, as indicated in Chapter 2, are the most popular and simple signal processing methods used to estimate the time delays in acoustic halls. In this study, the auto-correlation technique is used to estimate the bulk delays. The block diagram representing this estimation method is given in Figure 3.2. In the auto-correlation method, the echoed signal, $y(t)$, which is obtained after processing the input signal, $x(t)$, through the acoustic hall model, $H_h(z)$, is used to determine the bulk delay values.

In order to estimate the bulk delays of the acoustic hall model given in the previous section, a training signal, $x_T(t)$, which is a Gaussian white noise signal is used as the input signal. Using the auto-correlation technique, the echoed output

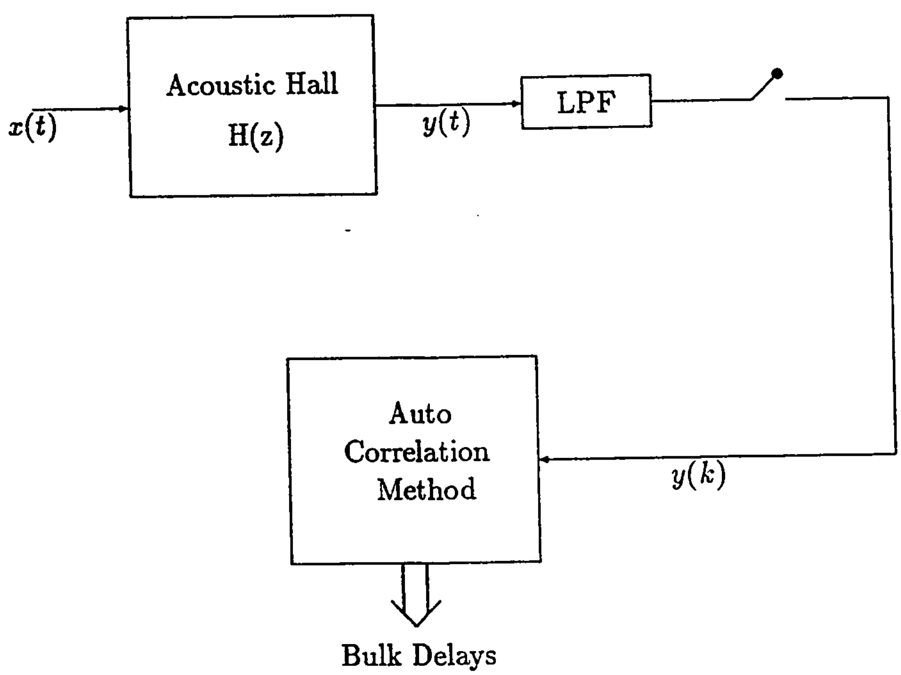


Figure 3.2: Auto-correlation technique to estimate the bulk delays.

signal is first low pass filtered by an FIR Low Pass Filter (LPF) where the cut of frequency, w_c, is 180 Hz. The output signal is low pass filtered since the bulk delay information is reserved at the low frequencies of an acoustic signal. The high frequencies are useful for the reverberation studies.

The filtered signal is then down sampled to get a simple and reasonable representation of the bulk delays of the acoustic hall model, $H_h(z)$. The analysis is then continued with the windowing. In this type of batch processing, it is suggested to use Hamming windowing since the basic operation is correlation. Thus, the sampled signal is weighted using a Hamming window. The bulk delay values of the hall is then estimated by the auto-correlation operation. In the auto-correlation method, the spectral density relationships of the filtered output signal are used to obtain estimates of the bulk delays.

In the time domain analysis, the bulk delay estimation is performed by the auto-correlation function

$$R_{yy}(k) = \sum_{i=0}^{M} y(i_1)y(i_2) \tag{3.5}$$

where the indices i_1 and i_2 are given as

$$i_2 = i_1 + d * k$$
$$i_1 = d * l$$
$$l = M - i$$
$$k = 1, 2, ..., N.$$

The value of M is determined with respect to the down sampling factor, d. For example, for a set of samples with 50000 points sampled at 8 kHZ, the value of M can be in a range of 3125 to 6250. This means that the down sampling is performed by a factor of 8 to 16. The value of N, however, determines the size of correlation function. The correlation function, $R_{yy}(k)$, is also referred as the auto-correlation vector and it is used to determine the estimates of the bulk delays. This estimates will be the estimated locations of the large bulk delays.

The correlation function in this case is in toeplitz structure which is solved to determine the estimates of the predictor and reflection coefficients as a linear combination of past values. Solution of the correlation function results in graphs representing the peak positions of the bulk delays of the acoustic hall model. In this study, these graphs are referred as the predictor coefficients plots. In order to determine accurate values of the largest K peaks of these plots, the predictor coefficients are normalized and those coefficients with magnitudes less than 0.25 are set to zero. This way the peak peaking algorithm is prevented to peak bias peak locations. The largest K peaks of the predictor coefficients are picked and titled as the bulk delays of the acoustic hall model, $H_h(z)$. The locations of these peaks are then stored and later transferred to the next step that is identification of the digital filter coefficients.

In the estimation process of such acoustic studies, a white noise signal is used as the input signal to estimate the initial bulk delay values. It is known that the impulse response of an acoustic hall varies for many reasons, such as the physical changes in the surrounding acoustic environment. Through out the equalization process, as it will be mentioned in the following sections, these bulk delays can be modified to accommodate for variations in the acoustic hall impulse response at certain times.

3.3.3 Filter Design

In the previous section, the bulk delays of the acoustic hall model are estimated. This has given the locations of the most dominant peaks in the impulse response; with these values performing the echo cancellation process will not be satisfactory because a limited number of values are available to represent the impulse response of the acoustic hall. Hence, it is required to identify the acoustic hall characteristics in more details for a proper echo cancellation.

Next step in designing a digital filter to model the acoustic hall is then identification of the required number of filter coefficients. From the acoustic hall model, $H_h(z)$, the following relationship can be given

$$y(z) = \frac{1}{1 + H(z)} x(z) \tag{3.6}$$

or alternatively

$$\frac{y(z)}{x(z)} = \frac{z^{-L_0}}{A(z)} \tag{3.7}$$

where

$$A(z) = 1 + \sum_{i=0}^{N} h_i(z) z^{-L_i} \tag{3.8}$$

$x(z)$ is the input noise signal and $y(z)$ is the output echoed signal. Referring to the original model, the input/output relationship in the finite difference form can be given by the time domain convolution equation as

$$y(t) = -\sum_{i=0}^{N} a_i y(t-i) + x(t-L_0) \tag{3.9}$$

where N is the filter order and a_is are the filter coefficients to be identified. Assuming an estimate of the output echoed signal $y(t)$ as

$$\hat{y}(t) = -\sum_{i=0}^{N} a_i y(t-i) \tag{3.10}$$

then the estimation error, $e(t)$, will be

$$e(t) = \hat{y}(t) - y(t) \tag{3.11}$$

or

$$e(t) = x(t-L_0) \tag{3.12}$$

which is a shift in the input noise signal by an amount of L_0. In order to reduce the effect of this error, the output echoed signal in the identification process is shifted by the same amount.

Due to the computational limitations, it is almost impossible to perform the echo cancellation process by identifying all the available filter coefficients in real time using a single commercial DSP board. For this reason, it is proposed to identify only those coefficients within the range of $[-I, I]$ of each bulk delay location. This is a reasonable assumption since these coefficients can be considered most effective in the intelligibility of speech signal in the public address systems. In determining the size of this range, that is the value of I, it is considered that each dominant peak location is separated by an approximate amount of 20 msec. Thus, the value of I can be varied between 5 to 20.

Estimate of the output echoed signal is considered in the identification process. Thus, the output echoed signal can be written as

$$y(t) = -\sum_{i=0}^{N} a_i y(t-i) \tag{3.13}$$

or, using an alternate notation for each sample point

$$y(n) = -\sum_{i=0}^{N} a_i y(n-i). \tag{3.14}$$

New form of the finite difference equation after incorporating the proposed assumption is given as

$$y(n) = -\sum_{k=0}^{K} \sum_{i=-I}^{I} h_k(i) y(n - \tilde{L}_k + i) \tag{3.15}$$

where $h_k(i)$s, also referred as the IIR filter coefficients of the acoustic hall, are the new coefficients to be identified, K is the number of bulk delays considered, and $\tilde{L}_k$ is the location of the kth bulk delay in the $\frac{1}{A(z)}$ polynomial. To simplify the notation, the location of each bulk delay is shifted I places backwards and the following equation is obtained

$$y(n) = -\sum_{k=0}^{K} \sum_{i=0}^{2I} h_k(i) y(n - L_k + i) \tag{3.16}$$

where L_k is the modified location of the kth bulk delay which is given as

$$L_k = \tilde{L}_k + I.$$

Identification of these coefficients is performed by solving a set of linear equations which is obtained by using the least-squares estimation method. This general approach to least-squares estimation of such models is also known as the covariance method. To form the set of linear equations, using this method, both sides of the above equation is multiplied by $y(n-(L_j+l))$ and the following normal equation is obtained

$$y(n)y(n-L_j-l) = -\sum_{k=0}^{K}\sum_{i=0}^{2I} h_k(i)y(n-L_k+i)y(n-L_j-l) \qquad (3.17)$$

where

$$j = 0,1,...,K$$
$$l = 0,1,...,2I$$

and L_j is location of the jth bulk delay which is also shifted I places backwards from the original jth bulk delay location. This equation can be put into a closed form like

$$\rho(L_j+l) = -\sum_{k=0}^{K}\sum_{i=0}^{2I} h_k(i)\rho(L_j+l-L_k+i) \qquad (3.18)$$

where

$$j = 0,1,...,K$$
$$l = 0,1,...,2I.$$

Then by letting

$$L_{jl} = L_j + l$$

the following can be written

$$\rho(L_{jl}) = -\sum_{k=0}^{K}\sum_{i=0}^{2I} h_k(i)\rho(L_{jl} - L_k + i) \tag{3.19}$$

where the set of $\{\rho(L_{jl})\}$ is the auto-correlation vector and the set of $\{\rho(L_{jl} - L_k + i)\}$ is the correlation matrix. They are also called the expected values of signals $y(n)$ and $y(n-(L_j+l))$, and $y(n-L_k+i)$ and $y(n-(L_j+l))$ respectively. Sizes of the auto-correlation vector and the correlation matrix are on order of a few hundreds. For instance, using the first 7 dominant bulk delays and an identification range of $[-8, 8]$, the size of auto-correlation vector is $(16+1)*7 = 119$ and the size of correlation matrix is $119 * 119$.

Now representing $\{\rho(L_{jl})\}$ by a vector like $\mathbf{B}(i)$ and $\{\rho(L_j + l - L_k + i)\}$ by a matrix like $\mathbf{R}(i,j)$ the following relationships can be written

$$\mathbf{B}(i) = E\{y(n)y(n-(L_{jj}+l))\} \tag{3.20}$$

$$\mathbf{R}(i,j) = E\{y(n-L_k+ii)y(n-(L_{jj}+l))\} \tag{3.21}$$

where

$i = (2I+1)*k+l-1$

$j = (2I+1)*k+ii-1$

$k = 0, 1, ..., K$

$l = 1, 2, ..., (2I+1)$

$ii = 1, 2, ..., (2I+1)$

$jj = 0, 1, ..., K.$

The final form of the normal equation is then

$$\mathbf{B}(i) = \sum_{j} \mathbf{h}(i)\mathbf{R}(i,j) \tag{3.22}$$

where

$\mathbf{B}(i) = \sum_{n=0}^{N} y(n)y(n + L_{jj} + l)$

$\mathbf{R}(i,j) = \sum_{n=0}^{N} y(n + L_k + ii)y(n + L_{jj} + l)$

$i = (2I + 1) * k + l - 1$

$j = (2I + 1) * k + ii - 1$

$k = 0, 1, ..., K$

$l = 1, 2, ..., (2I + 1)$

$ii = 1, 2, ..., (2I + 1)$

$jj = 0, 1, ..., K.$

$\mathbf{B}(i)$ is the auto-correlation vector of size $(2I + 1)K$, $\mathbf{R}(i,j)$ is the correlation matrix of size $(2I + 1)K * (2I + 1)K$, and $\mathbf{h}(i)$ is the IIR filter coefficients vector of size $(2I + 1)K$. Manual derivations showed that the correlation matrix is in a block symmetric form in which every block is in toeplitz matrix form. However, the overall matrix is not in a toeplitz form. Thus, available fast algorithms for solving toeplitz matrices are not adequate to use in this application for solving the normal equation. To solve this equation, however, other efficient methods are considered.

The identification process of the IIR filter coefficients starts with shifting each bulk delay location by an amount of I. It is recommended, in this study, to use the Cholesky decomposition for solving the above normal equation to get estimated values of the IIR filter coefficients. In the estimation process, the batch processing is preferred because in this kind of acoustic studies time for taking the corrective action is limited. In the acoustical echo cancellation studies, the equalizer filter coefficients are usually updated every 4-5 seconds. To compensate for the variations in impulse response of the large acoustic halls, the conventional recursive techniques are not proper to use. The batch processing techniques are faster than the sequential processing techniques. Therefore, in this study, batch processing is used to estimate the filter coefficients. The optimum values of these filter coefficients are stored for the use in designing the equalizer filter which is an adaptive FIR filter.

The whole acoustic hall filter modeling scheme, which is referred as the adaptation process, was shown in Figure 3.1. The adaptation process block, in the figure, can be described in more details as in the block diagram given in Figure 3.3. Since the objective is to identify or modify the filter coefficients for the equalization process, this process is referred as the adaptation process. In this study, the adaptation process is only performed at a very early stage in the acoustical echo cancellation system just to identify the IIR filter coefficients of the acoustic hall model. However, in the on-line implementations, these coefficients are needed to be updated at certain times to compensate for the rapid variations. An algorithm for modifying these coefficients will be described in the next chapter.

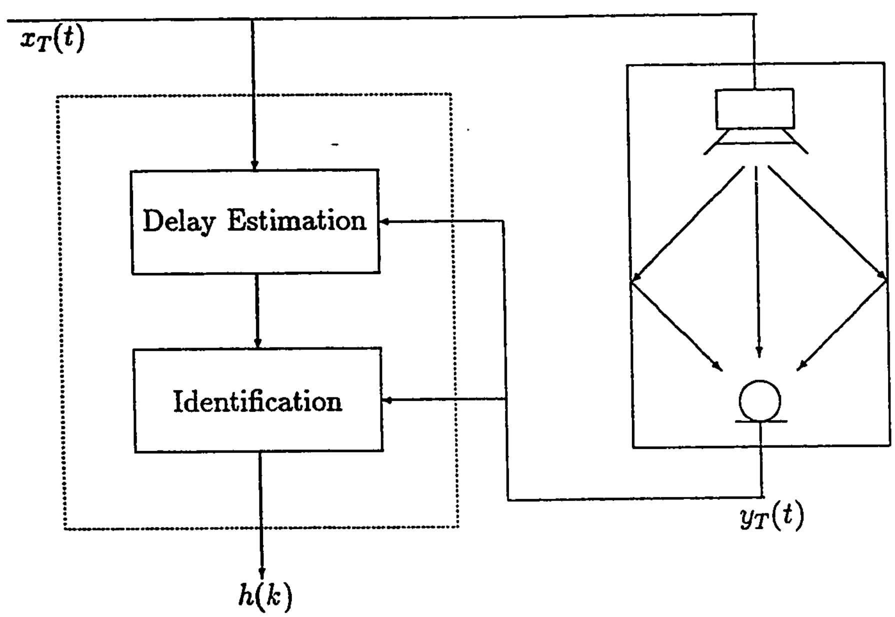


Figure 3.3: Block diagram of the adaptation process.

3.4 Acoustic Echo Equalizer

The second major step in designing an acoustical echo cancellation system for large halls is the equalization process. The equalization process basically consists of generating an echo replica which is constructed with use of the filter coefficients estimated in the previous section and the input speech signal. The echo replica is a modified form of the input speech signal which is obtained by multiplying the input speech signal with the estimated filter coefficients.

In this process, the acoustical echo cancellation is performed by subtracting the echo replica from the input speech signal. The signal obtained is then passed through the acoustic hall model or the real acoustic hall to analyze performance of the echo equalizer. The results obtained and the performance analysis are described in the simulations chapter. The equalizer filter design concept and structure of the acoustic echo equalizer are described in the following sections.

3.4.1 Equalizer Filter

The digital filter, which is used to represent characteristics of the acoustic hall, is assumed to be an IIR filter. An IIR digital filter model can be described as

$$H_h(z) = \frac{z^{-m_0}}{A(z)} \tag{3.23}$$

where

$$A(z) = 1 + \sum_{i=0}^{N} h_i(z) z^{-L_i}. \tag{3.24}$$

In this model, $h_i(z)$s are called the IIR filter coefficients used to describe characteristics of the echo path of the acoustic hall and N is the number of coefficients. The equalizer filter, in this case, is an FIR digital filter having the same coefficients. Model of this FIR digital filter can be given as

$$H_c(z) = B(z) \tag{3.25}$$

where

$$B(z) = 1 + \sum_{i=0}^{N} h_i(z) z^{-L_i} \tag{3.26}$$

which is the inverse of IIR filter and referred as the digital equalizer filter.

Coefficients of the equalizer filter, $h_i(z)$s, are estimated in the adaptation process. These coefficients are used in the cancellation of acoustic echoes. In the equalization process, these N filter coefficients are multiplied with the corresponding sample points of the input speech signal to produce the echo replica which is given as

$$\hat{y}(t) = \sum_{i=0}^{N} h_i x(t-i) \tag{3.27}$$

The size of N is usually very large and makes the equalization process difficult. However, it is proposed to limit the size of N to a few hundreds and perform as much multiplication as required. Samples of the input signal other than the multiplied ones are not modified; i.e. these values of the echo replica are assumed to be zero. This is a reasonable assumption since the input speech signal is needed to be modified at those frequencies where the effect of echo is destructive.

The equalizer filter model is assumed to process alone unless any variations take place in the acoustic hall environment. The complete structure of this filter is capable of handling such variations. The system designed for the acoustic echo cancellation can update coefficients of the equalizer filter to compensate for the changes in the impulse response of the acoustic echo path. Thus, the equalizer filter in a sense is an adaptive FIR filter. Several simulation studies are performed using this adaptive filter. The simulation results and the comparisons on performance of the adaptive filter are described in the next chapter.

3.4.2 Structure of the Acoustic Echo Equalizer

A block representation of the equalizer filter design scheme was given in Figure 3.1 as the equalization process. More descriptive form of this block is shown in Figure 3.4. Under the normal conditions, that is when the physical changes in the acoustic hall environment are not significant, the acoustic echo equalization is achieved through the following steps:

- The input speech signal, $x(t)$, is convolved by the filter coefficients, $h(k)$, to produce the echo replica, $\hat{y}(t)$.

- The echo replica obtained is then subtracted from the input speech signal as shown in the figure.

- The signal formed after subtraction is passed through the acoustic hall. This signal is supposed to construct a sound field in the acoustic hall without any echo.

This process continues as far as the quality of perception or the speech intelligibility in the acoustic hall is acceptable. These conditions are monitored in the

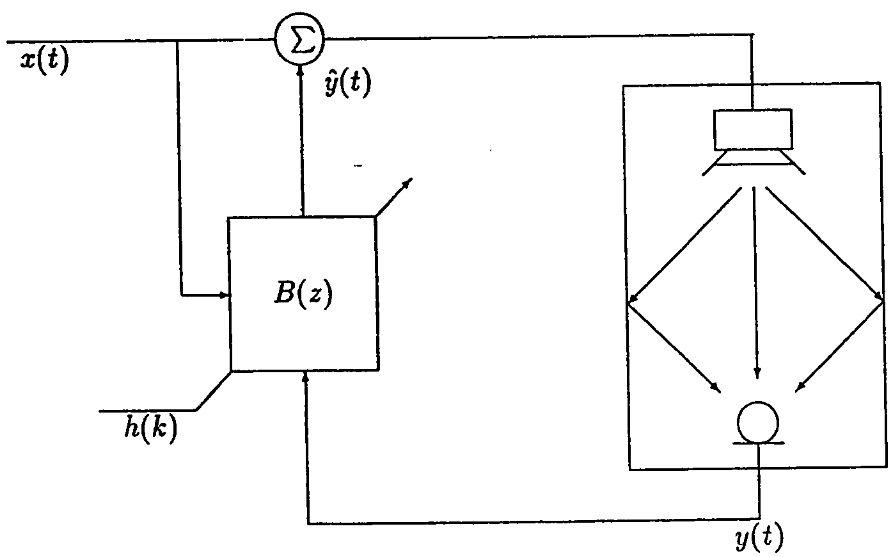


Figure 3.4: Block diagram of the cancellation process.

equalization process with use of the output signal $y(t)$. At any time, if the speech intelligibility is reduced below the acceptable limits, the filter coefficients, $h(k)$, are updated through the adaptation process. Modification of the equalizer filter coefficients can be performed at the background, like when there is no speech message to be passed through the acoustic hall. In the public address systems, transfer of the speech message is not as frequent as in any other acoustic system. Therefore, modifying the equalizer filter coefficients does not necessarily affect the acoustic echo cancellation process.

3.5 Advantages and Limitations of the Proposed System

In practice, an acoustical echo canceller operates on the overall spectrum of the impulse response. It deals with minimum a few thousands of tap coefficients which require several DSP boards to run parallel for the equalization process. In the straightforward FIR equalizer filter implementations, an acoustical echo canceller requires more than 4000 taps and results in a complex hardware.

On the other hand, the subband echo cancellers can decrease the total computational requirement by splitting the signal into several frequency bands and providing a separate echo canceller for each band. However, the effect of residual echo at the band gap limits the number of bands to 4 or 8 bands. An echo cancellation system having such subband structure as well requires several DSP boards to cover an echo duration of a few hundreds of msecs.

The proposed echo cancellation system can be implemented in real time on a single DSP board or even on a 486 machine equipped with A/D - D/A utilities. It can handle acoustic hall system which has an echo duration of as much as several hundreds of msecs. Number of tap coefficients to be identified in this system is not more than 150 coefficients. Identifying and also updating such a less number of coefficients do not require complex computations. Since the computational complexity is reduced considerably, the identification and modification schemes are chosen to be the basic and simple techniques. However, more enhanced algorithms and techniques can also be adapted to the proposed echo cancellation system.

In the real time equalization process, the proposed cancellation system can behave better than the commercially existing acoustical echo cancellation systems. It does not require a complex hardware for the real time implementations and operates only on a few number of tap coefficients. Updating a few number of tap coefficients requires less iterations and results in a faster convergence speed. Analysis on the initial convergence is given in the simulations chapter.

One disadvantage of the proposed echo cancellation system is that if the echo information exist at the high frequencies of the input speech signal then the performance of the cancellation system reduces sharply. This is because the equalizer filter assumes that the echo is available at the low frequencies and processes at these frequencies. Moreover, the echo generated by the acoustic coupling cannot be cancelled by the proposed system. In this study, the acoustic coupling between the microphone and the loudspeaker system is not considered.

3.6 Summary

The objective of an acoustic echo cancellation study is to find a way to make the convergence speed fast and reduce the computational complexity. This chapter presented a new system design on the cancellation of acoustical echoes of large halls. The proposed cancellation system assumes that the acoustic hall is given as an IIR filter model. When the hall model is an IIR digital filter then the most appropriate filter structure for the equalization process is the FIR digital filter structure.

An adaptive FIR filter designed for cancelling the acoustic echoes in the public address systems. The equalizer filter is designed in a simple structure by which the computational effort is reduced considerably. The application of identifying reduced number of tap coefficients for cancellation acoustical echoes reduces the computational complexity significantly. Therefore, a real-time implementation of the proposed system on a single DSP board is possible. Simulations studies are presented in the next chapter. The results are used to verify the performance of the proposed echo cancellation system.

Chapter 4

SIMULATION RESULTS AND PERFORMANCE EVALUATION

4.1 Introduction

The objective of this simulation study is to validate the performance of the proposed echo cancellation scheme. The performance of the echo cancellation system is demonstrated graphically by using the cross-correlation method.

A brief summary of the off-line simulation tasks can be listed as

1 - Passing the white noise signal, $x(t)$, through an acoustic hall model

2 - Sensing the output echoed signal, $y(t)$

3 - Low pass filtering the output echoed signal

4 - Down sampling the filtered form of the output echoed noise signal

5 - Windowing the down sampled signal using Hamming weighting

6 - Correlating the windowed signal to form the auto-correlation vector, $ryy(k)$

7 - Solving the toeplitz system to determine the predictor and reflection coefficients, $a(k)$ and $rc(k)$ respectively

8 - Normalizing the predictor coefficients and setting those $a(k)$s that are less than 0.25 to zero

9 - Picking the maximum K absolute values of the updated predictor coefficients and labeling their locations as bulk delays

10 - Forming the proposed form of the normal equation

11 - Solving this normal equation to determine the equalizer filter coefficients

12 - Constructing the echo replica signal, $\hat{y}(t)$, and subtracting it from the input speech signal

13 - Passing the signal obtained after subtraction through the same acoustic hall model.

Two different acoustic hall models are used in evaluating the proposed echo cancellation scheme. First acoustic hall model is chosen as AR model where a single pole is used to simulate the acoustic hall characteristics. However, the second model is chosen to be an all-zero filter that is a MA model. In the following sections, these hall models are described in details and the performance of both the acoustic hall filter and the equalizer filter are discussed.

4.2 System Used in the Simulations

The off line simulations are performed using the software programmed for the proposed echo cancellation system. This software is written in C language and its flowcharts and listings are provided in the Appendices. The principal operation of this software is reading a noise and a reference speech signals and then generating an equalized output speech signal. Steps of this process are listed in the previous section.

The test setup used for the simulation studies is shown in Figure 4.1. The reference input signal (noise or speech) is recorded and digitized through the Chimera signal processing board's Serial Voice Interface (SVI) unit. This digitized signal is then stored in a data file using the Hear utility of the Chimera. The software coded to simulate the echo canceller then reads the noise and speech files stored in data

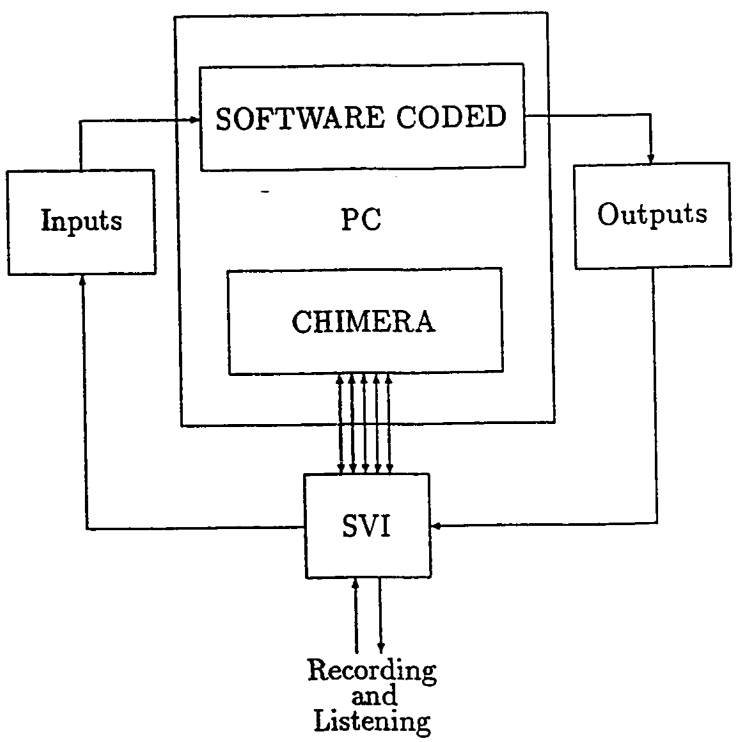


Figure 4.1: Off line simulation system.

files and generates the equalized speech data file. This equalized data file is then listened again using the Hear utility of the Chimera board via SVI to test the performance of the proposed system from the human perception point of view.

4.3 Acoustic Echo Path Characteristics

The acoustic echo path models used in the simulations are derived as the reverberation models given in Chapter 1. An autoregressive model and a moving average model are used in simulating the behavior of a large acoustic hall. These two models are also tested with real speech signals to confirm their operations from human ear perception point of view.

The first model, which is basically an IIR filter with a low pass configuration, is chosen to have a single pole for simplicity and for testing the proposed cancellation algorithm. Transfer function of this IIR filter model is derived as

$$H(z) = \frac{z^{-m_0}}{(1-\alpha z^{-1})(1-\beta z^{-m_1})} \tag{4.1}$$

where m_0 is used to simulate the ITD and m_1 is taken to be the delay time. In generating the echoed signal, the following difference equation is used

$$y(n) = \alpha y(n+1) + \beta y(n+m_1) - \alpha\beta y(n+m_1+1) \tag{4.2}$$

$$y(n) = y(n) + x(n+m_0) \tag{4.3}$$

where $y(n)$ is the sampled form of the output echoed signal, $x(n)$ is that of the input source signal, and α and β are the feedback coefficients. The sampled input and output signals are stored backwards. Therefore, the + signs are used instead of – signs in the difference equations. Since the input and output signals are stored backwards throughout the simulation studies performed, the delay times are added to the sample index in computing $y(n)$.

The second model, however, is chosen to be a more complicated model so that to verify the performance of the proposed echo cancellation scheme. This second model is used at the next step in simulation studies after confirming proper operation of the algorithm. In this part of simulation study, the acoustic hall considered is modeled by an FIR filter. The transfer function of the second model is derived as follows

$$H(z) = z^{-L_0}(1 + \sum_{i=1}^{3} \alpha_i z^{-L_i}) \tag{4.4}$$

where again L_0 is used for the ITD and L_is are taken to be the delay times. The difference equation for this FIR configuration is given as

$$y(n) = x(n + L_0) \tag{4.5}$$

$$y(n) = y(n) + \alpha_i x(n + L_0 + L_i) \tag{4.6}$$

for $i = 1, 2, 3$.

Proper variables for both AR and MA models are chosen to efficiently represent the typical large acoustical hall responses [15]. α and β decaying variables are used for the attenuation of major echo signals. Whereas, m_0, m_1, L_0, and L_i time delay variables are used to generate repeatations of the echo paths. In this research study, since the aim is to study the affect of the early reflections, variables concerned with reverberation region of acoustical halls are not included in the derived acoustical hall models.

4.4 Performance of the Acoustic Hall Filter

In evaluating the performance of the acoustic hall filter models given in the previous section, human speech signals are tested through these models. However, a white noise signal is preferred in the graphical evaluation, since it is not adequate to use speech signals. Due to the natural repeatations in speech signals, it is difficult to determine the position and magnitude of possible echo signals. Thus, to represent the performance of these acoustic hall models graphically, a white Gaussian noise signal is used in the simulation studies.

The white Gaussian noise signal is first passed through the acoustical hall model derived to obtain the echoed output signal. This output signal is filtered with an LPF where cut off frequency is 180 kHz. Later, the filtered output signal is down sampled by a factor of 16 and windowed using a Hamming window.

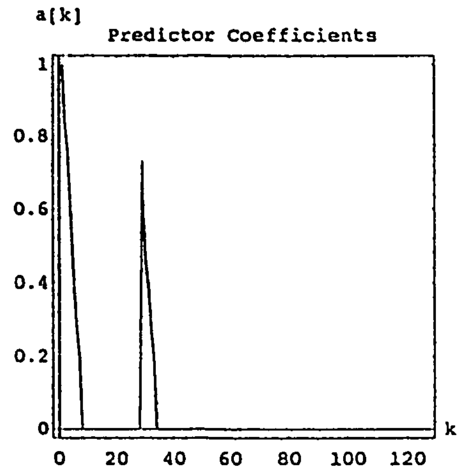
Down sampled and windowed output echo signal is than used to form the auto-correlation vector which is obtained for deriving the predictor. To derive the predictor, the linear prediction method is used. The predictor coefficients are determined by solving the toeplitz structure obtained by the auto-correlation method. The toeplitz structure is solved to find the partial correlation or the reflection coefficients and the predictor coefficients.

The simulation studies are performed with 3 sec. noise signal recorded at 8 kHz. This signal is stored using 24000 sample points. In the IIR acoustical hall model, the value of m_0 is set to 256 and the value of m_1 is set to 450. This means that the direct sound is assumed to reach the listeners location after 32 msec. However, the first echo signal assumed in this model reaches the listeners location 56 msec. after the direct sound is heard.

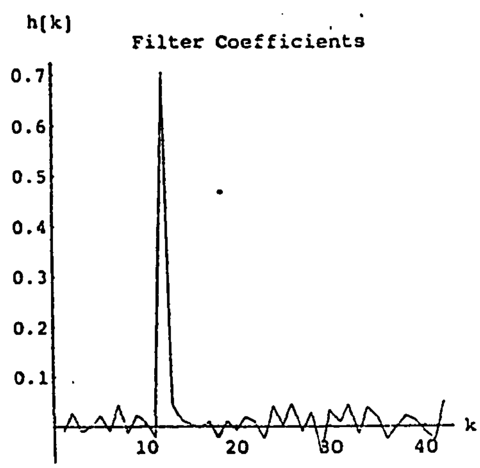
Auto-correlation vector is formed only up to a duration of 250 msec. since this is the region which has the maximum affect on the perception of speech signals. After solving the toeplitz structure obtained using this vector, the predictor coefficients are plotted.

In the next step of the proposed scheme, K most dominant peaks of the predictor coefficients are peaked and titled as the bulk delays. A guard region (or the search area) is considered in peaking the dominant bulk delays. The size of this region is determined from the down sampling factor and from the equalization region proposed for the echo cancellation scheme. This equalization region is also called the regional equalizer filter size. The value of K and the size of the equalizer filter are varied to test the performance of peak peaking algorithm. The output from the bulk delay estimation algorithm is given in Figure 4.2. Plot of the predictor coefficients, given in Figure 4.2-a, shows a clear isolated bulk delay value at the location 30x16 = 450 which is the value of m_1.

In the FIR acoustical hall model, however, the initial time delay value, L_0, is set to 320 (representing 40 msec.) and L_is are set to 640, 890, and 1530 respectively. As seen from the plot, the bulk delay estimation scheme used in this study is promising in determining the location of large time delay. Moreover, the graphs



(a)



(b)

Figure 4.2: Predictor and equalizer filter coefficients - IIR model.

obtained throughout the simulation study showed that the two acoustic hall models derived are giving clear echo signals at the defined positions. Both IIR and FIR models are used in this study for simulating the characteristics of large acoustic halls so that to study the performance of the proposed echo cancellation scheme from different perspectives.

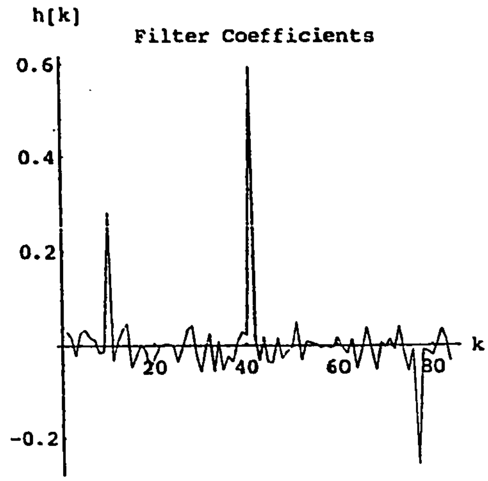
4.5 Performance of the Equalizer

Structure of the equalizer filter is determined from the number of bulk delays used and the size of the equalization region considered. In the formation of this filter, the down sampling factor is also taken into consideration.

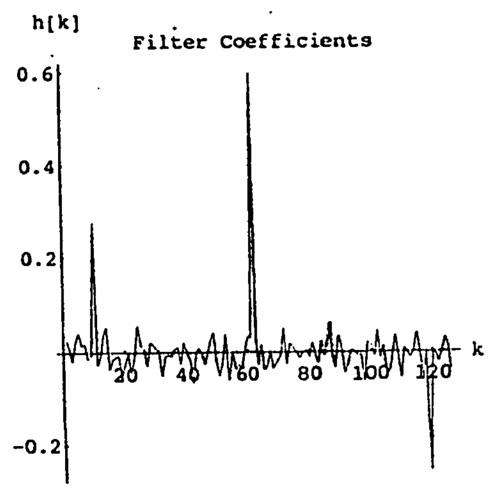
After determining the positions of the bulk delays, only those coefficients around the equalization region of each bulk delay location are identified. These identified coefficients are referred as the equalizer filter coefficients. The number of these coefficients is determined from the size of the equalization region and number of valid bulk delays. For example, using 2 bulk delays and an equalization region size of [-10,10], the number of these coefficients will be 42. This means that only 42 taps are used in the cancellation process rather than a few thousands. The proposed scheme considers that the size of tap coefficients vector can be varied with the bulk delay values and the size of equalization region. Graphical representation of tap coefficients determined for 4 bulk delay values with an equalization region of [-10,10] is given in Figure 4.2-b for the IIR model. The plot of filter coefficients shows only 42 values since the bulk delay estimation algorithm is resulted with two valid bulk delays locations.

The simulations results for the FIR acoustical hall model are given in Figure 4.3. In this figure, equalizer filter coefficients obtained for the model are given. The plot, given in Figure 4.3-a, shows the filter coefficients obtained when the valid bulk delay value is 4 and the regional equalizer filter size is 10. The second plot, given in Figure 4.3-b, shows the same result with 6 bulk delay values. Comparing these two plots, it can be concluded that having more valid bulk delay values and different equalization region size do not change the tap coefficients. However, the size of the regional filter will certainly have greater affect when a real acoustic hall is considered. In the real acoustic halls, it is not possible to find a single isolated peak as a bulk delay. The peaks in such cases happen in clusters due to the dispersion of sound waves.

The performance of the equalizer filter is also evaluated graphically. To do this the input noise signal and the echoed output signal are first cross-correlated to see the positions and the magnitudes of the direct sound and the reflections. In the second step of the performance evaluation, the input noise signal and the equalized output signal are cross-correlated. These cross-correlation values show the level of cancellation by using the proposed scheme. The plots of the cross-correlations obtained for the IIR model are given in Figures 4.4. The cross-correlation result between the input speech signal and the echoed output speech signal is given in Figure 4.4-a. However, Figure 4.4-b shows the cross-correlation result obtained after the cancellation between the input speech signal and the equalized speech signal. As shown from the plots the proposed echo cancellation system is behaved

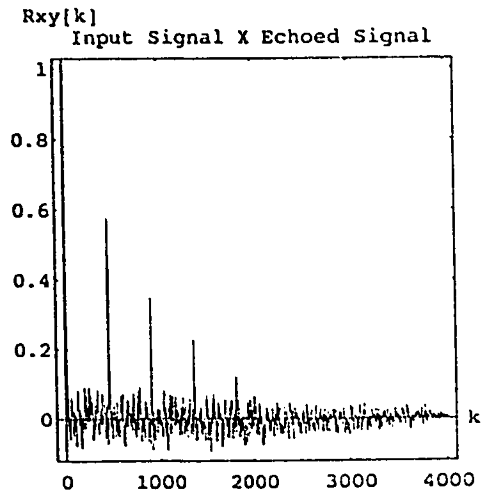


(a)

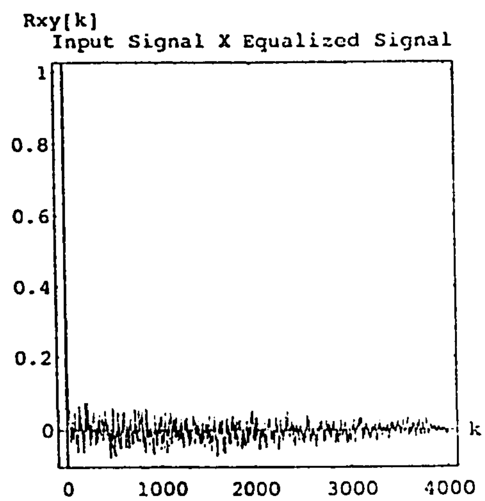


(b)

Figure 4.3: Equalizer filter coefficients - FIR model, for 4 and 6 valid bulk delay values.



(a)



(b)

Figure 4.4: Cross-correlation results for the IIR model.

23

very efficient in cancelling the echo signals generated from the IIR acoustic hall model. In Figure 4.5 and 4.6, the cross-correlations obtained for the FIR acoustic hall model are given. The result of cross-correlating the input noise signal and the output echoed noise signal is shown in Figure 4.5. However, the results of cross-correlating the input noise signal and the equalized noise signal are shown Figure 4.6. The first plot, given in Figure 4.6-a, is obtained with 4 bulk delay values and the equalization region size of [-10,10]. The plot in Figure 4.6-b, however, is obtained with 6 bulk delay values and the same equalization region.

Comparing the figures of the cross-correlation results, it can be concluded that using only 2 bulk delays with a region size of [-5,5] (which requires only 22 coefficients) a good performance is obtained for the IIR model derived for the simulations. However, for the FIR model it is required to use 4 bulk delays and a minimum regional equalizer filter size of 10 for an acceptable echo cancellation. The size of regional equalizer filter must be increased for the real time echo cancellation in real acoustic halls. The proposed echo cancellation system can also be adapted to such changes in the regional equalizer filter size.

For a possible real time implementation necessary operations count needed is approximately calculated for the derived IIR acoustic hall model. It is assumed that the optimal values of the valid bulk delays and the regional equalizer filter size are used in computing the computational complexity. Therefore, for an adaptive FIR digital filter of order 11 only 22 additions/multiplications per 125 msec. are needed for the real time equalization. Approximate values of the number of operations per

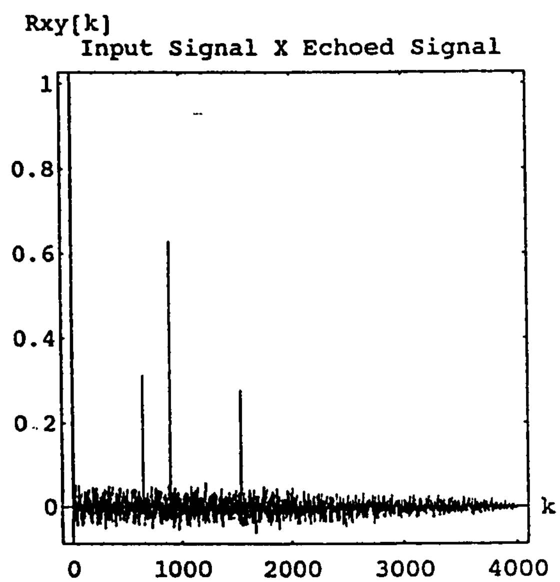
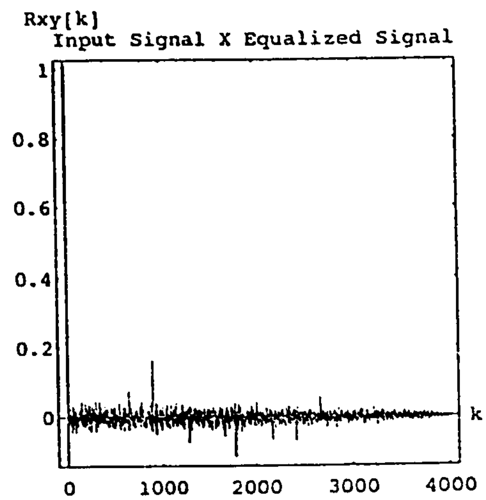
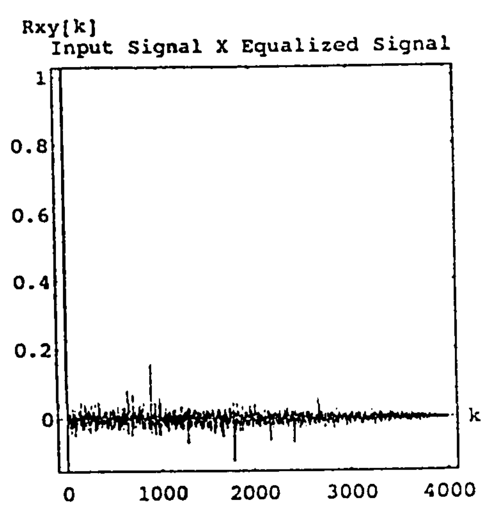


Figure 4.5: Cross-correlating input and echoed signals - FIR model.



(a)



(b)

Figure 4.6: Cross-correlating input and equalized signals - FIR model.

second required for the estimation, identification, and equalization processes are listed in Table 4.1. The proposed system can easily be implemented for a real time processing using a 486 PC equipped with analog/digital (A/D and D/A) utilities since the total number of operations per second needed is about 0.7 millions. On the other hand, the proposed system does not require a large amount of memory. The software program used in the simulation studies requires approximately 500 Kbyte RAM. Indeed, for a real time implementation the proposed system will require lesser amount of RAM since it will not be necessary to include an acoustic hall model and related codings in the software.

4.6 Summary

The proposed echo cancellation system is tested in this chapter using two acoustical hall models. The performance of the proposed algorithm is presented graphically for the estimation of the bulk delays and for the cancellation of acoustical echoes. It has been seen that the proposed scheme can take the proper action for a single isolated echo and for a number of separate echoes. Optimum values of the number of bulk delay values and the equalization region sizes for both IIR and FIR acoustic hall models used in the simulation studies are reported.

Table 4.1: Operations count required for a real time implementation.

Type of Process	No. of Operations/Sec.
Estimation	
Filtering	180000
Sampling	24000
Windowing	48000
Toeplitz	15000
Peaking peaks	25000
Others	100000
Identification	
Normal equation	120000
Cholesky	30000
Others	100000
Equalization	
Cancellation	18000
Others	30000
TOTAL	690000

Chapter 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The study performed in this thesis involves designing an echo cancellation system for large acoustic halls. This system is designed to operate on the early reflections exist in a sound field of a large acoustic hall. The design procedure consists of estimating the large time delays of the acoustic hall and identifying proper equalizer coefficients for an efficient echo cancellation process.

The designed echo cancellation scheme is tested through off-line simulation studies for different acoustical environments that are derived mathematically. Different acoustical echo path models are considered for these off-line simulations. The performance of the derived acoustic hall models are tested graphically and as well by human perception point of views.

A new technique is proposed for the formation of tap coefficients of the acoustic echo canceller. A complete simulation study is performed using this new technique to determine the optimum values of the equalization variables for the derived acoustic hall models. It has been reported that this new method can sharply reduce the number of tap coefficients involved with the acoustical echo cancellation. This way the computational complexity can also be reduced considerable and this will enable any possible implementation on a single DSP board.

The performance evaluation of echo canceller with this new technique is demonstrated graphically and tested with human speech signals. It is concluded the proposed echo cancellation system is promising for the acoustic echo cancellation with lesser computations and flexible usage.

5.2 Recommendations for Future Studies

Throughout this research study several interesting further research areas have been identified. A good collection of these research areas are summarized below:

- Real time implementations of the proposed acoustical echo cancellation scheme on the commercially available DSP boards.
- Optimization of the cancellation algorithm for faster and flexible implementations in the real time environment.

- Testing the proposed system through more complicated acoustical hall models and using different digital filter structures for modeling.

- Studying the affect of including the reverberation region in the acoustical hall models derived for the off-line simulations.

- Investigating and testing other estimation techniques for large time delay estimation and other faster adaptation techniques for identifying equalizer filter coefficients.

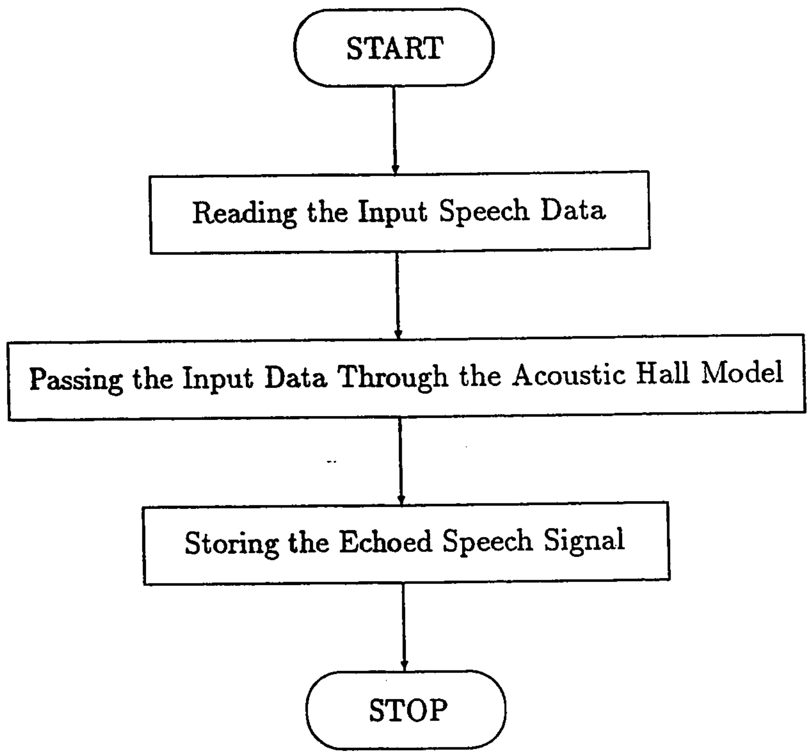
Appendix A

FLOWCHARTS OF THE SOFTWARE PROGRAMS

This Appendix illustrates the software programs used in the off-line simulations with the flowcharts. It is divided into two main sections, namely the echo generation software and the echo cancellation software. The echo cancellation software is also divided into three subsections for the simplicity of illustration. Subsections of the echo cancellation software are

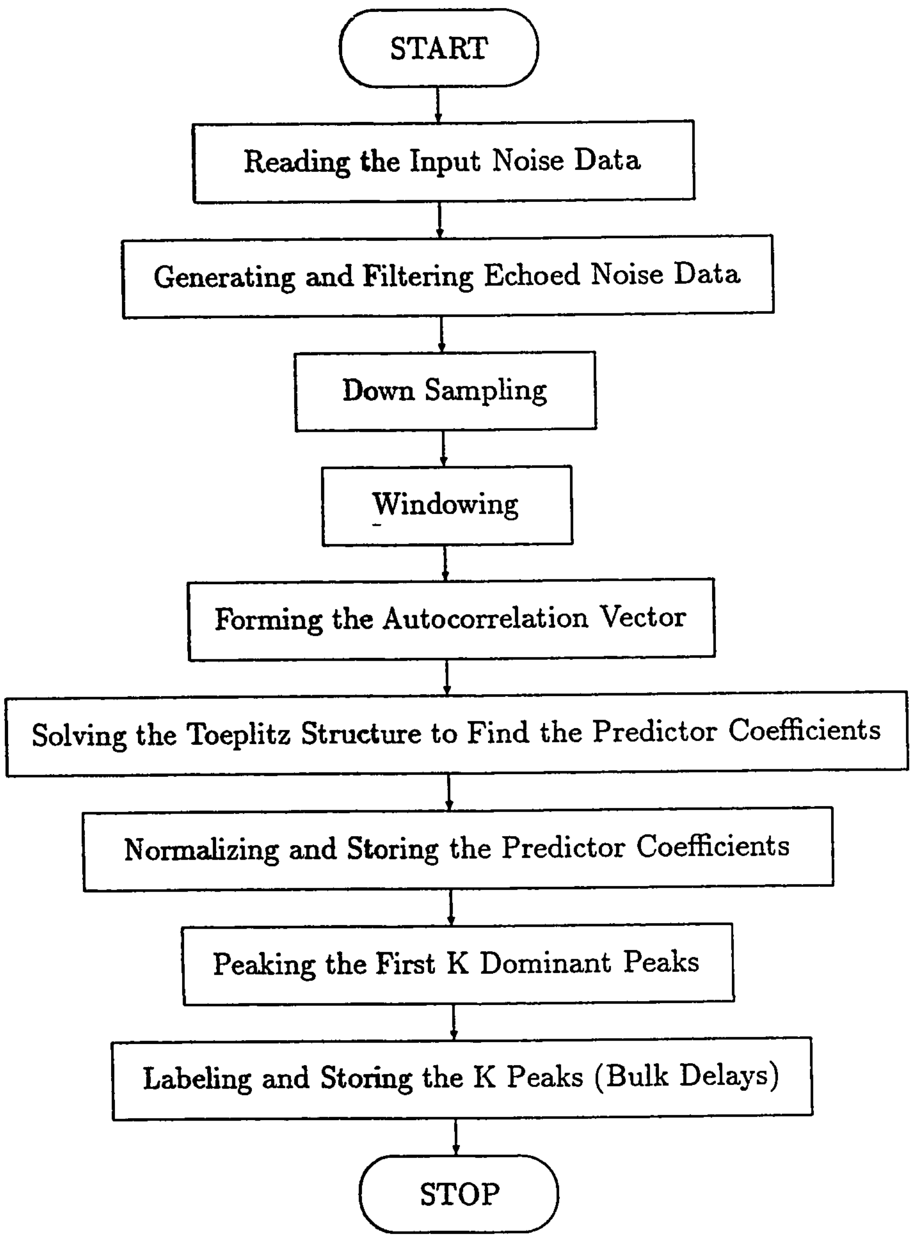
- 1. Bulk delay estimation part
- 2. Filter coefficients identification part
- 3. Echo cancellation part.

FLOWCHART OF THE ECHO GENERATION SOFTWARE

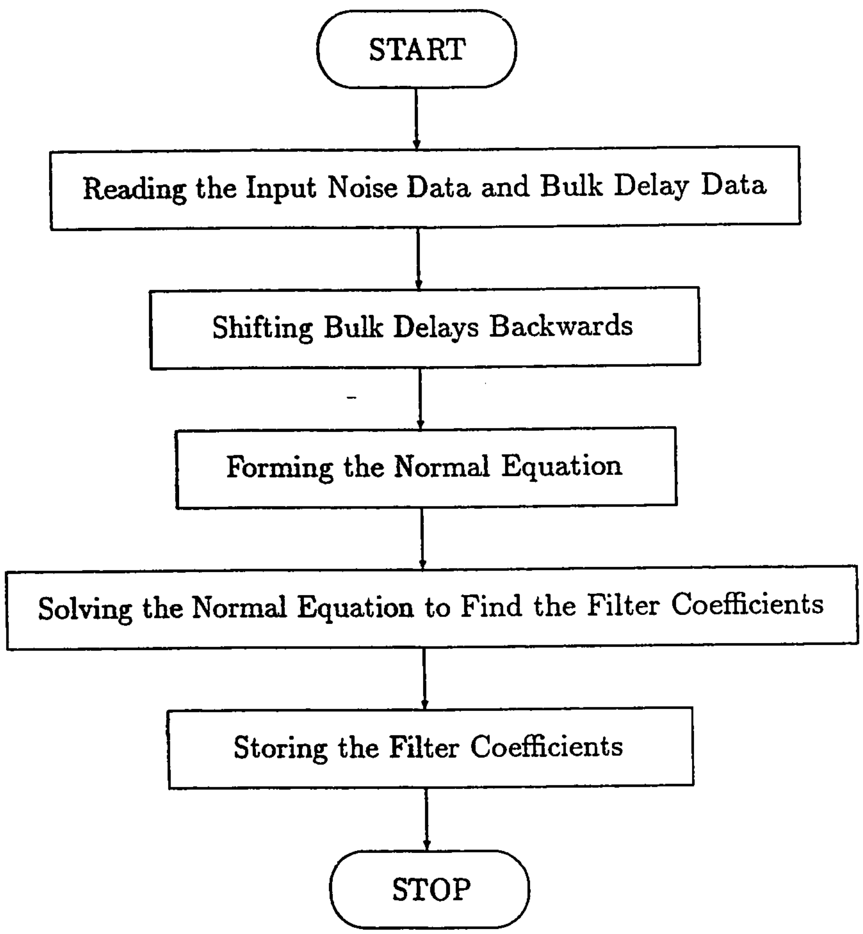


23

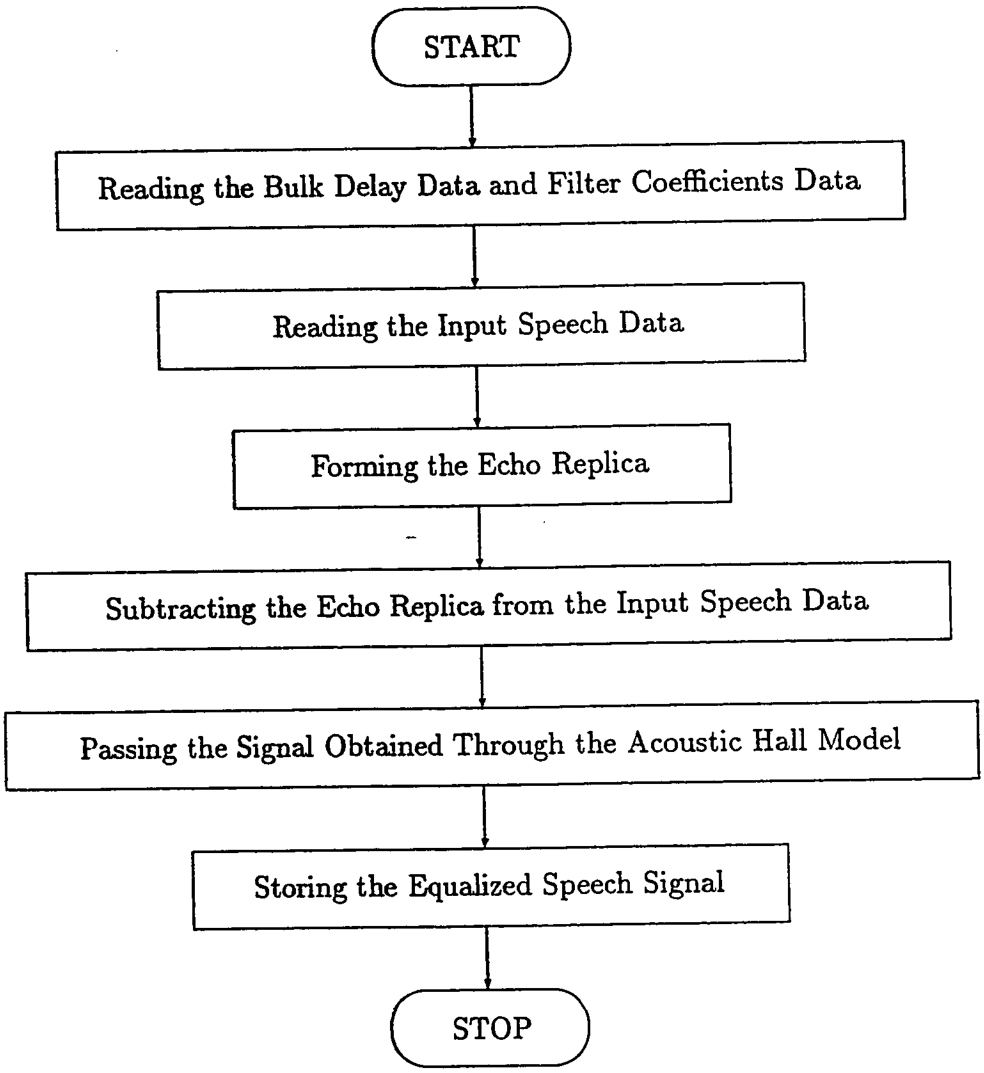
FLOWCHART OF THE BULK DELAY ESTIMATION PART



FLOWCHART OF THE FILTER COEFFICIENTS IDENTIFICATION PART



FLOWCHART OF THE ECHO CANCELLATION PART



Appendix B

LISTINGS OF THE SOFTWARE PROGRAMS

Listings of the software programs are given in this Appendix. In each program listing, a short introduction part is included for the explanation of input/output files used in that specific program. Listings of the subroutines used throughout these programs are given at the end of this Appendix.

ECHO GENERATION PROGRAM

```
/**********************************************************************/
/* ECHO GENERATION PROGRAM                                            */
/*     Inputs  = Speech or Noise Signal Data                          */
/*     Outputs = Echoed Speech or Noise Signal                        */
/**********************************************************************/
#include      <fcntl.h>
#include      <sys\types.h>
#include      <sys\stat.h>
#include      <io.h>
#include      <stdio.h>
#include      <float.h>
#include      <math.h>
#include      <graph.h>
#define       NS            24000
FILE          *read_file;
FILE          *read1_file;
FILE          *write_file;
unsigned int count;
int           ck,numread,nact=0,buffer[400],syn_bfr[NS],l0=320,m0=256;
int           m1=450;lz[3]={640,890,1530};
float         alp=0.8,beta=0.7,huge y[NS],huge s[NS],s_1(),y_1();
float         alpha[3]={0.36,0.8,0.45};
char          ch,choice[10],inputfile[40],outfile[40],tmp[NS];
/**********************************************************************/
/*** MAIN PROGRAM                                                   ***/
/**********************************************************************/
main() {int quit=0;
 while(!quit)
```

```
{printf("\n\n                    ECHO GENERATION PROGRAM \n\n");
 printf("\n\n\n         ********** MAIN SELECTION MENU **********");
 printf("\n\n\n\n\n      1......Generate ECHO file              \n");
 printf("\n\n            2......Exit                            \n");
 printf("\n\n\n          ****************************************");
 printf("\n\n\n\n\n           >>> Enter your choice. >>> : ");
 gets(choice); ch=atoi(choice);
 switch(ch) {case  1: g_echo(); break; default: quit=1;}}}
/*********************************************************************/
/*** END OF MAIN PROGRAM                                          ***/
/*********************************************************************/
/*********************************************************************/
/*** ECHO GENERATION :                                            ***/
/*********************************************************************/
int g_echo() {unsigned int ib, k, n1; int i, nn, rcode=1; float y0;
repeat: printf("\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n ");
 printf("\n\n  *****************************************************");
 printf("\n\n  ***          ACOUSTIC HALL MODEL                 ***");
 printf("\n\n  ***                                              ***");
 printf("\n\n  ***     1.     IIR Filter Model                  ***");
 printf("\n\n  ***     2.     FIR Filter Model                  ***");
 printf("\n\n  *****************************************************");
 printf("\n\n\n\n          >>> Enter your choice. >>> : ");
 gets(choice); ch=atoi(choice);
 switch(ch) {case 1: ck=1; goto start; case 2: ck=2; goto start;
            default: goto repeat;}
start: printf("\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n ");
 printf("******** INITIALIZE THE FILES ******** \n\n\n");
 printf("\n   >>>  Enter INPUT file name (.dat)   >>> "); gets(inputfile);
```

```
printf("\n   >>>  Enter OUTPUT file name (.ech)  >>> "); gets(outfile);
if((read_file=fopen(inputfile,"r+b"))==NULL)
 {printf("\n ERROR! ....... COULD NOT OPEN INPUT FILE  \n");
  printf("\n\n   >>>  Press ENTER to continue ....   >>> ");
  gets(choice); exit();}
if((write_file=fopen(outfile,"w+b"))==NULL)
 {printf("\n ERROR! ....... COULD NOT OPEN OUTPUT FILE  \n");
  printf("\n\n   >>>  Press ENTER to continue ....   >>> ");
  gets(choice); exit();}
/**********************************************************************/
/* Reading the Input Speech Data                                      */
/**********************************************************************/
 numread=fread((char *)buffer,sizeof(int),64,read_file);
 for(i=0;i<numread;i++) syn_bfr[i]=buffer[i]; rcode=1; nn=64;
 while(rcode==1) {numread=fread((char *)buffer,sizeof(int),400,read_file);
   for(i=0;i<numread;i++) {ib=NS-nn-i-1; s[ib]=(float)buffer[i];}
   nn=nn+numread; if(ib>=NS) rcode=0;}
 syn_bfr[4]=-8000; syn_bfr[5]=0; count=64;
/**********************************************************************/
/* Passing the Input Data Through the Acoustic Hall Model             */
/**********************************************************************/
 for(k=64;k<NS;k++)
  {ib=NS-k-1; if(ck==1) gg1(ib); if(ck==2) gg2(ib); y0=y[ib];
   if(y0>=32768.0) y0=32767.0; if(y0<=-32768.0) y0=-32767.0;
   syn_bfr[count]=(int)(y0); count++;}
/**********************************************************************/
/* Storing the Echoed Speech Signal                                   */
/**********************************************************************/
 n1=fwrite((char *)syn_bfr,sizeof(int),count,write_file); nact=nact+n1;
```

```
printf("\n\n   >>> Number of stored samples = %u",nact);
printf("\n\n   >>> Press ENTER to continue ....   >>> ");
gets(choice); fcloseall();}
/*********************************************************************/
/*** END OF ECHO GENERATION                                        ***/
/*********************************************************************/
```

ECHO CANCELLATION PROGRAM

```
/*********************************************************************/
/* ECHO CANCELLATION PROGRAM                                         */
/*    Inputs  = Training Noise Signal Data                           */
/*              Reference Speech Signal Data                         */
/*    Outputs = Reflection Coefficients                              */
/*              Predictor Coefficients                               */
/*              Bulk Delays                                          */
/*              Auto-correlation Vector                              */
/*              Correlation Matrix                                   */
/*              Equalizer Filter Coefficients                        */
/*              Equalized Speech Signal                              */
/*********************************************************************/
#include <fcntl.h>
#include <sys\types.h>
#include <sys\stat.h>
#include <io.h>
#include <stdio.h>
#include <float.h>
#include <math.h>
#include <graph.h>
#define  NS    24000        /* Total number of samples 3 sec.        */
#define  DS    1500         /* Down sampling to DS samples           */
```

```
#define  DF    16            /* Down sampling factor                    */
#define  AV    125           /* Size of the autocorrelation vector      */
#define  FS    10            /* Regional equalizer filter size          */
#define  SA    (2*FS)/DF+1   /* Search area for peaking peaks           */
#define  IR    (NS/150)/DF   /* Initial region to be skipped            */
#define  FM    3             /* Number of FIR filter multipliers        */
#define  BD    4             /* Number of bulk delay considered         */
FILE *read1_file;
FILE *read2_file;
FILE *read3_file;
FILE *read4_file;
FILE *write1_file;
FILE *write2_file;
FILE *write3_file;
FILE *write4_file;
int m0=256,m1=450,ipk[AV+1],bulk[BD+1],lres[175],buffer[400],ck=0,nd;
int l0=320;chk,syn_bfr[NS],lz[FM]={640,890,1530};
float alpa=0.8,beta=0.7,hf[101],xd[DS+1],ry[AV+1],rc[AV+1],a[AV+1];
float b[175],r[175][175],coef[175],sres[175],x[NS],y[NS],s[NS],x_1();
float apk[AV+1];y_1();s_1(),alpha[FM]={0.36,0.8,0.45};
char ch,choice[2],inputfile[12],outfile[12],tmp[NS];
/******************************************************************/
/*** MAIN PROGRAM                                               ***/
/******************************************************************/
main() {int quit=0; while(! quit)
  {printf("\n\n\n\n\n\n\n\n\n\n                                         ");
   printf("\n *************** MAIN SELECTION MENU ***************    ");
   printf("\n\n ***       1.     Find DELAYS                       *** ");
   printf("\n\n ***       2.     Find COEFFICIENTS                 *** ");
```

```
printf("\n\n ***        3.      Perform CANCELLATION                *** ");
printf("\n\n ***        4.      Exit                                *** ");
printf("\n\n ***************************************************** ");
printf("\n\n\n\n            >>> Enter your choice. >>> : ");
gets(choice); ch=atoi(choice);
switch(ch) {case 1: delay(); break; case 2: filter(); break; case 3:
          cancel(); break; default: quit=1;}}}
/*******************************************************************/
/*** END OF MAIN PROGRAM                                         ***/
/*******************************************************************/
/*******************************************************************/
/*** BULK DELAY ESTIMATION PART :                                ***/
/*******************************************************************/
delay() {unsigned int is,i; int i1,j,k,n,m,imax,imin;
 float y1,x1,x2,x3,alp,ss,at,pre_peak,peak,max;
repeat: printf("\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n ");
 printf("\n\n *****************************************************");
 printf("\n\n ***            ACOUSTIC HALL MODEL               ***");
 printf("\n\n ***                                              ***");
 printf("\n\n ***        1.      IIR Filter Model              ***");
 printf("\n\n ***        2.      FIR Filter Model              ***");
 printf("\n\n *****************************************************");
 printf("\n\n\n\n            >>> Enter your choice. >>> : ");
 gets(choice); ch=atoi(choice);
 switch(ch) {case 1: chk=1; goto start; case 2: chk=2; goto start;
          default: goto repeat;}
start: printf("\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n ");
 printf("******** BULK DELAY PROGRAM ******** \n\n\n");
/********************************************************************/
```

```
/* Reading the Input White Noise Data                                    */
/************************************************************************/
 if((read1_file=fopen("white.gas","r+b"))==NULL)
  {printf("\n ERROR...! Could not open");
   printf("\n >>> Press ENTER to continue. >>> "); exit();}
 printf("\n\n >>> Enter OUTPUT file name (a.dat)    >>> ");
 gets(outfile); if((write1_file=fopen(outfile,"w"))==NULL)
  {printf("\n ERROR...! Could not open");
   printf("\n >>> Press ENTER to continue. >>> "); exit();}
 printf("\n\n >>> Enter OUTPUT file name (rc.dat)   >>> ");
 gets(outfile); if((write2_file=fopen(outfile,"w"))==NULL)
  {printf("\n ERROR...! Could not open");
   printf("\n >>> Press ENTER to continue. >>> "); exit();}
 printf("\n\n >>> Enter OUTPUT file name (bulk.dat) >>> ");
 gets(outfile); if((write3_file=fopen(outfile,"w"))==NULL)
  {printf("\n ERROR...! Could not open");
   printf("\n >>> Press ENTER to continue. >>> "); exit();}
 fread((char *)tmp,sizeof(char),NS,read1_file); is=NS;
 while(is>0) {s[is]=(float)tmp[is]; is--;} f_design();
/************************************************************************/
/* Generating & Filtering Echoed Noise Data                              */
/************************************************************************/
 for(i=0;i<NS;i++)
  {is=NS-i-1; if(chk==1) gecho1(is); if(chk==2) gecho2(is); y1=0.0;
   for(n=0;n<=100;n++) y1=y1+hf[n]*y_1(is+n); x[is]=y1;}
/************************************************************************/
/* Down Sampling                                                         */
/************************************************************************/
 for(i=0;i<DS;i++) {is=DS-i; xd[is]=x_1(DF*is);}
```

```
/*******************************************************************/
/* Windowing                                                       */
/*******************************************************************/
 x1=(float)DS; for(i=0;i<DS;i++)
  {x2=(float)i; x3=0.54-0.46*cos(x2*6.28318/x1); xd[i]=x3*xd[i];}
/*******************************************************************/
/* Forming the Autocorrelation Vector                              */
/*******************************************************************/
 for(m=0;m<=AV;m++)
  {ry[m]=0.0; n=DS-m-1; for(i=0;i<=n;i++) ry[m]=ry[m]+xd[i]*xd[i+m];}
 for(m=1;m<=AV;m++) ry[m]=ry[m]/ry[0]; ry[0]=1.0;
/*******************************************************************/
/* Solving the Toeplitz Structure to Find Predictor Coefficients   */
/*******************************************************************/
 rc[0]=-ry[1]/ry[0]; a[0]=1; a[1]=rc[0]; alp=ry[0]+ry[1]*rc[0];
 for(m=1;m<AV;m++)
  {ss=0.0; for(i=0;i<=m;i++) ss=ss+ry[m+1-i]*a[i]; rc[m]=-ss/alp;
   n=(m+1)/2; for(i=1;i<=n;i++)
    {is=m+1-i; at=a[i]+rc[m]*a[is]; a[is]=a[is]+rc[m]*a[i]; a[i]=at;}
   a[m+1]=rc[m]; alp=alp+rc[m]*ss; if(alp<=0) break;}
/*******************************************************************/
/* Normalizing & Storing Predictor Coefficients                    */
/*******************************************************************/
 max=0.0; for(i=IR+1;i<=AV;i++) if(fabs(a[i])>max) max=fabs(a[i]);
 for(i=0;i<=AV;i++) {ipk[i]=i; a[i]=fabs(a[i]); if(a[i]<=0.25*max) a[i]=0.0;
   fprintf(write1_file,"\n %e",a[i]); fprintf(write2_file,"\n %e",rc[i]);}
/*******************************************************************/
/* Peaking the First K Dominant Peaks                              */
/*******************************************************************/
```

```
k=0; pre_peak=0.0; for(i=IR+1;i<=AV;i++)
 {imax=i+SA; imin=i-SA; if(imin<0) {imin=0; imax=i+2*SA;} i=imax;
  if(imax>=AV) imax=AV-1; peak=a[imin]; m=imin;
  for(j=imin;j<=imax;j++) {if(a[j]>peak) {peak=a[j]; m=j;}}
  if(peak!=pre_peak)
   {if(peak>0.0) {apk[k]=peak; ipk[k]=m; zero(ipk[k]); k++;}}
  pre_peak=peak;}
 for(i=0;i<k;i++) {for(j=i;j<k;j++) {if(apk[i]<apk[j]) {x1=apk[i];
  apk[i]=apk[j]; apk[j]=x1; i1=ipk[i]; ipk[i]=ipk[j]; ipk[j]=i1;}}}
 if(k>BD) k=BD;
 for(i=0;i<k;i++){for(j=i;j<k;j++) {if(ipk[i]>ipk[j]) {x1=apk[i];
  apk[i]=apk[j]; apk[j]=x1; i1=ipk[i]; ipk[i]=ipk[j]; ipk[j]=i1;}}}
/*********************************************************************/
/* Labeling and Storing the K Peaks (Bulk Delays)                    */
/*********************************************************************/
 for(i=0;i<k;i++)
  {bulk[i]=DF*ipk[i]; fprintf(write3_file,"\n %d",bulk[i]);}
 printf("\n\n\n\n >>> Press ENTER to continue....      >>> ");
 gets(choice); fcloseall(); nd=k;}
/*********************************************************************/
/*** END OF BULK DELAY ESTIMATION                                  ***/
/*********************************************************************/

/*********************************************************************/
/*** FILTER COEFFICIENTS IDENTIFICATION PART :                     ***/
/*********************************************************************/
filter() {unsigned int is,i; int j,k,l,kk,ll,m; float y0;
 printf("\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n ");
 printf("******** EQUALIZER FILTER PROGRAM ******** \n\n\n");
/*********************************************************************/
```

```
/* Reading the Input Noise Data & Bulk Delay Data                         */
/***********************************************************************/
 if((read1_file=fopen("white.gas","r+b"))==NULL)
  {printf("\n ERROR...! Could not open");
   printf("\n >>> Press ENTER to continue. >>> "); exit();}
 printf("\n\n >>> Enter INPUT file name (bulk.dat)  >>> ");
 gets(inputfile); if((read2_file=fopen(inputfile,"r"))==NULL)
  {printf("\n ERROR...! Could not open");
   printf("\n >>> Press ENTER to continue. >>> "); exit();}
 printf("\n\n >>> Enter OUTPUT file name (b.dat)    >>> ");
 gets(outfile); if((write1_file=fopen(outfile,"w"))==NULL)
  {printf("\n ERROR...! Could not open");
   printf("\n >>> Press ENTER to continue. >>> "); exit();}
 printf("\n\n >>> Enter OUTPUT file name (r.dat)   >>> ");
 gets(outfile); if((write2_file=fopen(outfile,"w"))==NULL)
  {printf("\n ERROR...! Could not open");
   printf("\n >>> Press ENTER to continue. >>> "); exit();}
 printf("\n\n >>> Enter OUTPUT file name (coef.plo) >>> ");
 gets(outfile); if((write3_file=fopen(outfile,"w"))==NULL)
  {printf("\n ERROR...! Could not open");
   printf("\n >>> Press ENTER to continue. >>> "); exit();}
 printf("\n\n >>> Enter OUTPUT file name (coef.dat) >>> ");
 gets(outfile); if((write4_file=fopen(outfile,"w"))==NULL)
  {printf("\n ERROR...! Could not open");
   printf("\n >>> Press ENTER to continue. >>> "); exit();}
 fread((char *)tmp,sizeof(char),NS,read1_file);
 is=NS; while(is>0) {s[is]=(float)tmp[is]; is--;} for(i=0;i<NS;i++)
  {is=NS-i-1; if(chk==1) gecho1(is); if(chk==2) gecho2(is);}
 for(i=1;i<=nd;i++) fscanf(read2_file,"\n %d",&bulk[i]);
```

```
/********************************************************************/
/* Shifting the Bulk Delays Backwards                              */
/********************************************************************/
 for(i=1;i<=nd;i++) bulk[i]=bulk[i]-FS;
/********************************************************************/
/* Forming the Normal Equation i.e. B = R * COEF                   */
/********************************************************************/
 k=0; l=1;
loop: i=(2*FS+1)*k+l-1; b[i]=0.0;
 for(j=0;j<=1000;j++) b[i]=b[i]+y_1(j)*y_1(j+bulk[k+1]+l);
 fprintf(write1_file,"\n %e",fabs(b[i])); kk=0; ll=1;
loop1: j=(2*FS+1)*kk+ll-1; r[i][j]=0.0;
 for(m=0;m<=1000;m++)
  r[i][j]=r[i][j]+y_1(m+bulk[k+1]+l)*y_1(m+bulk[kk+1]+ll);
 fprintf(write2_file,"\n %e",fabs(r[i][j])); ll++; if(ll>(2*FS+1))
  {ll=1; kk++; if(kk>=nd) goto down;} goto loop1;
down: fprintf(write2_file,"\n "); l++;
 if(l>(2*FS+1)) {l=1; k++; if(k>=nd) goto skip;} goto loop;
/********************************************************************/
/* Solving the Normal Equation to Find the Filter Coefficients     */
/********************************************************************/
 skip: result();
/********************************************************************/
/* Storing the Filter Coefficients                                 */
/********************************************************************/
 for(i=0;i<nd*(2*FS+1);i++) {fprintf(write3_file,"\n %e",fabs(coef[i]));
   fprintf(write4_file,"\n %e",coef[i]);}
 printf("\n\n\n\n >>> Press ENTER to continue....      >>> ");
 gets(choice); fcloseall();}
```

23

```
/*********************************************************************/
/*** END OF FILTER COEFFICIENTS IDENTIFICATION                     ***/
/*********************************************************************/

/*********************************************************************/
/*** ECHO CANCELLATION PART :                                      ***/
/*********************************************************************/
cancel() {unsigned int k,is,n,m; int i,l,rcode; float y0;
 printf("\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n\n ");
 printf("******** CANCELLATION PROGRAM ******** \n\n\n"); ck=1;
 printf("\n\n >>> Enter INPUT file name (bulk.dat)  >>> ");
 gets(inputfile); if((read1_file=fopen(inputfile,"r"))==NULL)
  {printf("\n ERROR...! Could not open");
   printf("\n >>> Press ENTER to continue. >>> "); exit();}
 printf("\n\n >>> Enter INPUT file name (coef.dat)  >>> ");
 gets(inputfile); if((read2_file=fopen(inputfile,"r"))==NULL)
  {printf("\n ERROR...! Could not open");
   printf("\n >>> Press ENTER to continue. >>> "); exit();}
 printf("\n\n >>> Enter INPUT file name (spec.dat)  >>> ");
 gets(inputfile); if((read3_file= fopen(inputfile,"r+b"))==NULL )
  {printf("\n ERROR...! Could not open");
   printf("\n >>> Press ENTER to continue. >>> "); exit();}
 printf("\n\n >>> Enter OUTPUT file name (spec.equ) >>> ");
 gets(outfile); if((write1_file=fopen(outfile,"w+b"))==NULL)
  {printf("\n ERROR...! Could not open");
   printf("\n >>> Press ENTER to continue. >>> "); exit();}
 /*********************************************************************/
 /* Reading the Bulk Delay Data & Filter Coefficients Data            */
 /*********************************************************************/
  for(i=1;i<=nd;i++) fscanf(read1_file,"\n %d",&bulk[i]);
```

```
for(i=0;i<nd*(2*FS+1);i++) fscanf(read2_file,"\n %e",&coef[i]);
/*********************************************************************/
/* Reading the Input Speech Data                                     */
/*********************************************************************/
 k=fread((char *)buffer,sizeof(int),64,read3_file);
 for(i=0;i<k;i++) syn_bfr[i]=buffer[i]; syn_bfr[4]=-8000; syn_bfr[5]=0;
 rcode=1; n=64;
 while(rcode==1) {k=fread((char *)buffer,sizeof(int),400,read3_file);
  for(i=0;i<k;i++) {is=NS-n-i-1; s[is]=(float)buffer[i];}
  n=n+k; if(is>=NS) rcode=0;}
/*********************************************************************/
/* Forming the Echo Replica                                          */
/*********************************************************************/
 y0=0.0; n=0; m=64; for(k=64;k<NS;k++) {is=NS-k-1;
  for(i=1;i<=nd;i++) {for(l=1;l<=2*FS+1;l++)
   {y0=y0+coef[n]*s_1(is+bulk[i]+l-FS); n++;}}
/*********************************************************************/
/* Subtracting the Echo Replica from the Input Speech Data           */
/*********************************************************************/
   x[is]=s[is]-y0;
/*********************************************************************/
/* Passing the Signal Obtained Through the Acoustic Hall Model       */
/*********************************************************************/
   if(chk==1) gecho1(is); if(chk==2) gecho2(is);
   syn_bfr[m]=(int)(y_1(is)/8.0); n=0; y0=0.0; m++;}
/*********************************************************************/
/* Storing the Equalized Speech Signal                               */
/*********************************************************************/
 fwrite((char *)syn_bfr,sizeof(int),m,write1_file);
```

```
printf("\n\n\n\n >>> Press ENTER to continue....      >>> ");
gets(choice); ck=0; fcloseall();}
/***********************************************************************/
/*** END OF ECHO CANCELLATION                                        ***/
/***********************************************************************/
```

FUNCTIONS AND SUBROUTINES

```
/**********************************************************************/
/* Echo Generator - IIR Model. Used for echo generation.              */
/**********************************************************************/
int gg1(kk) unsigned int kk;
{if(kk>=NS) return;
 y[kk]=alp*y_1(kk+1)+beta*y_1(kk+m1)-alp*beta*y_1(kk+m1+1);
 y[kk]=y[kk]+s_1(kk+m0);}
/**********************************************************************/
/* Echo Generator - FIR Model. Used for echo generation.              */
/**********************************************************************/
int gg2(kk) unsigned int kk;
{int ti; if(kk>=NS) return; y[kk]=s_1(kk+10);
 for(ti=0;ti<3;ti++) y[kk]=y[kk]+alpha[ti]*s_1(kk+lz[ti]+10);}
/**********************************************************************/
/* Echo Generator - IIR Model. Used for echo cancellation.            */
/**********************************************************************/
int gecho1(t) unsigned int t;
{if(t>=NS) return;
 y[t]=alpa*y_1(t+1)+beta*y_1(t+m1)-alpa*beta*y_1(t+m1+1);
 if(ck==0) y[t]=y[t]+s_1(t+m0); else y[t]=y[t]+x_1(t+m0);}
/**********************************************************************/
/* Echo Generator - FIR Model. Used for echo cancellation.            */
/**********************************************************************/
```

```
int gecho2(t) unsigned int t;
{int ti; if(t>=NS) return;
 if(ck==0) {y[t]=s_1(t+10);
  for(ti=0;ti<FM;ti++) y[t]=y[t]+alpha[ti]*s_1(t+lz[ti]+10);}
 else {y[t]=x_1(t+10);
   for(ti=0;ti<FM;ti++) y[t]=y[t]+alpha[ti]*x_1(t+lz[ti]+10);}}
/**********************************************************************/
/* Low Pass Filter (LPF) with Wc = 180 Hz                             */
/**********************************************************************/
int f_design()
{int fi,fj; float f1,f2,pi=3.1415927; for(fi=0;fi<=100;fi++)
  {f2=0.54-0.46*cos(2.0*pi*fi/100.0); f1=0.1414*(fi-50);
   if(f1==0.0) {f1=1.0; goto step;} f1=sin(f1)/f1;
step: f1=(0.1414/pi)*f1*f2; hf[fi]=f1;}}
/**********************************************************************/
/* Subroutine to Solve the Normal Equation                            */
/**********************************************************************/
int result()
{int i,j,k,lk,yeter; float smax,rmax,rval,xmult,sum; yeter=nd*(2*FS+1);
 for(i=0;i<yeter;i++) {lres[i]=i; smax=0.0; for(j=0;j<yeter;j++)
   if(abs(r[i][j])>smax) smax=abs(r[i][j]); sres[i]=smax;}
 for(k=0;k<yeter-1;k++) {rmax=0.0; for(i=k;i<yeter;i++)
   {rval=abs(r[lres[i]][k])/sres[lres[i]]; if(rval<=rmax) goto loop;
   j=i; rmax=rval;
 loop:;}
    lk=lres[j]; lres[j]=lres[k]; lres[k]=lk; for(i=k+1;i<yeter;i++)
    {xmult=r[lres[i]][k]/r[lk][k]; for(j=k+1;j<yeter;j++)
     r[lres[i]][j]=r[lres[i]][j]-xmult*r[lk][j]; r[lres[i]][k]=xmult;}}
  for(k=0;k<yeter-1;k++) {for(i=k+1;i<yeter;i++)
```

```
 b[lres[i]]=b[lres[i]]-r[lres[i]][k]*b[lres[k]];}
coef[yeter-1]=b[lres[yeter-1]]/r[lres[yeter-1]][yeter-1];
for(i=yeter-2;i>=0;i--) {sum=b[lres[i]]; for(j=i+1;j<yeter;j++)
 sum=sum-r[lres[i]][j]*coef[j]; coef[i]=sum/r[lres[i]][i];}}
/*******************************************************************/
/* Other External Functions Used                                  */
/*******************************************************************/
float s_1(kk) unsigned int kk; {if(kk>=NS) return(0); return(s[kk]);}

float y_1(kk) unsigned int kk; {if(kk>=NS) return(0); return(y[kk]);}

float x_1(kk) unsigned int kk; {if(kk>=NS) return(0); return(x[kk]);}

int zero(k) int k; {int i; for(i=1;i<=SA;i++)
 {if(k<AV-SA) {a[k-i]=0.0; a[k+i]=0.0;}}}
```

Appendix C

NOMENCLATURE

d	down sampling factor
$E(.)$	expectation operator
g	feedback gain
$h(k)$	filter coefficients vector
L_0	initial delay line
L_i	reflected delay lines
$\tilde{L}_k$	exact delay location
L_k	modified delay location
m_0	initial time delay
m_{a_i}, m_{b_j}	feedback coefficients
n	time index with respect to t
$\mathbf{R}$	vector or matrix notations
RT_{60}	reverberation time
$R_{xy}(t)$	cross-correlation function
$R_{yy}(t)$	auto-correlation function

S	surface area
t	time symbol
V	volume
$x_T(t)$	training signal
$\hat{y}(t)$	estimate of $y(t)$
α, β	feedback coefficients
$\bar{\alpha}$	average absorption coefficient
μ	convergence speed constant
$\rho(.)$	auto-correlation operator
τ	delay line length
$*$	convolution operator
$\{.\}$	set of values
$[.]^T$	transpose of [.]

Appendix D

GLOSSARY

Alignment : Any mechanical or electronic adjustment to a system to bring it into conformance with some standard set of values.

Ambient Noise : The prevailing sound field in a room in the absence of an applied signal from a loudspeaker or a sound source.

Attenuation : A decrease in signal amplitude from one point to another, or the process causing this decrease.

Bandwidth : The difference between the lower and the upper cutoff frequencies.

Comb Filter : Any device that produces a series of sharp attenuation notches in the output response that resemble the teeth of a comb.

Crosstalk : Undesired energy appearing in one signal path as a result of coupling from other signal paths.

Cutoff Frequency : The frequency that is identified with the transition between a pass band and an adjacent attenuation band of a system.

Delay : The time between the arrival of a direct sound and a discrete echo of that sound.

Digital Equalizer : An equalizer whose frequency modification circuits operate in the digital domain.

Direct Path : The single straight line path between a sound source and a listener or a sensing microphone.

Direct Sound : The sound wave that travels along the direct path, known as the shortest distance between the source and the listener.

Distortion : The distortion of a signal is an undesired change in the waveform of that signal.

Early Reflections : The first few reflected sounds that arrive at the listener's location, usually containing a single reflection in each path.

Echo : A wave, reflected or returned with sufficient magnitude and delay, to be perceived as a wave distinct from that directly transmitted.

Echo Cluster : A sequence of closely spaced echoes, usually preceded and followed by a delay longer than that between the echoes in the cluster.

Equalization : The process of modifying the frequency response of an audio signal.

Equalizer : A device designed to compensate for an undesired amplitude-frequency or phase-frequency characteristic, or both, of a system.

Filter : Any network that attenuates a portion of the audio frequency spectrum.

Level : The magnitude of a quantity considered in relation to a reference value.

Mixer : A device having two or more adjustable inputs and a common output, which operates to combine linearly the separate input signals (in a desired proportion) to produce an output signal.

Pass Band : The range of frequencies that are attenuated by less then 3 dB.

Power Amplifier : An amplifier which drives a utilization device such as a mixer or a loudspeaker.

Reflection : The return of a waveform from an obstruction in its path.

Reverberant Field : The sound field in an area in which the sound source itself is no longer the strongest component.

Reverberation : A sequence of echoes that are closely spaced in time that it is impossible to distinguish one from another.

Reverberation Time : When a steady-state sound source is stopped, the time it takes for the sound pressure level to decrease by 60 dB.

Signal-to-Noise Ratio : The ratio of the amplitude of a signal to the amplitude of the noise measured at the same point in the system.

Sound Field : The complex combination of the direct and reflected energy that exists in a listening environment.

White Noise : A random noise containing equal energy in each 1-Hz band within the audio spectrum.

Bibliography

[1] Sondhi, M.M., and W. Kellermann, *"Adaptive Echo Cancellation for Speech Signals"*, Bell Laboratories, Murray Hill, N.J., 1990.

[2] Murano, K., S. Unagami, and F. Amano, *"Echo Cancellation and Applications"*, IEEE Comm. Magazine, PP. 49-55, January 1990.

[3] Beranek, L.L., *"Concert Hall Acoustics - 1992"*, J. of Acous. Soc. of Ame., PP. 1-39, July 1992.

[4] Davis, D., Davis, C., *Sound System Engineering*, Mac Millan, Inc., 1987.

[5] Schroeder, M.R., *"New Method of Measuring Reverberation Time"*, J. of Acous. Soc. of Ame., Vol. 37, PP. 409-412, 1965.

[6] Kawakami, F., and K. Yamaguchi, *"Space-ensemble Average of Reverberation Decay Curves"*, J. of Acous. Soc. of Ame., Vol. 70, PP. 1071-1082, 1981.

[7] Kellermann, W., *"Analysis and Design of Multirate Systems for Cancellation of Acoustical Echoes"*, IEEE ICASSP, PP. 2570-2573, 1988.

[8] Makino, S., and Y. Kaneda, *"Acoustic Echo Canceller Algorithm Based on the Variation Characteristics of a Room Impulse Response"*, IEEE ICASSP, PP. 1133-1136, 1990.

[9] Sikorav, J., *"Experiments to Identify and Track Non Stationarities in Audio Conference Room"*, IEEE ICASSP, PP. 2566-2569, 1988.

[10] Nelson, P.A., H. Hamada, and S.J. Elliott, *"Adaptive Inverse Filters for Stereophonic Sound Reproduction"*, IEEE Trans. on Sign. Procs., PP. 1621-1632, July 1992.

[11] Gay, S.L., and R.J. Mammone, *"Fast Converging Subband Acoustic Echo Cancellation Using RAP on the WE DSP16A"*, IEEE ICASSP, PP. 1141-1144, 1990.

[12] Chen, J., H. Bes, J. Vandewalle, and P. Janssens, *"A New Structure for Sub-band Acoustic Echo Canceler"* IEEE ICASSP, PP. 2574-2577, 1988.

[13] Hatty, B., *"Recursive Least Squares Algorithms Using Multirate Systems for Cancellation of Acoustical Echoes"*, IEEE ICASSP, PP. 1145-1148, 1990.

[14] Davis, D., *"The LEDE Concept"*, Audio, PP. 48-58, August 1987.

[15] Oppenheim, A.V. (ed.), *Applications for Digital Signal Processing*, Prentice Hall, Inc., Englewood Cliffs, N.J., Chapter 2, 1978.

[16] Schroeder, M.R., D. Gottlieb, and K.F. Siebrasse, *"Comparative Study of European Concert Halls: Correlation of Subjective Preference with Geometric and Acoustic Parameters"*, J. of Acous. Soc. of Ame., Vol. 56, PP. 1195ff, 1974.

[17] Dentino, M., J. McCool, and B. Widrow, *"Adaptive Filtering in the Frequency Domain"*, Proc. of IEEE, PP. 1658-1659, December 1978.

[18] Ferrara, E.R., *"Fast Implementation of LMS Adaptive Filters"*, IEEE Trans. on Acous. Spe. and Sign. Procs., PP. 474-475, August 1980.

[19] Haykin, S., *Adaptive Filter Theory*, Prentice Hall, Inc., Englewood Cliffs, N.J., 1986.

[20] Yasukawa, H., S. Shimada, and I. Furukawa, *"Acoustic Echo Canceller with High Speed Quality"*, IEEE ICASSP, PP. 49.8.1-49.8.4, 1987.

[21] Ching, P.C., Y.T. Chan, and K.C. Ho, *"Constrained Adaptation for Time Delay Estimation with Multipath Propagation"*, IEE Proc. F, PP. 453-458, October 1991.

[22] Kirsteins, I.P., *"High Resolution Time Delay Estimation"*, IEEE ICASSP, PP. 451-454, 1987.

[23] Pallas, M., and G. Jourdain, *"Active High Resolution Time Delay Estimation for Large BT Signals"*, IEEE Trans. on Sign. Procs., PP. 781-788, April 1991.

[24] Lee, H., J.M. Silkaitis, and D.P. Sullivan, *"Optimal Partial-Discrete Algorithm for Time Delay Estimation"*, Spe. Procs., PP. 265-272, 1989.

[25] Parsons, T., *Voice and Speech Processing*, McGraw Hill, Inc., 1987.

Vita

İlke Levent AKIN

Born at Adıyaman, TÜRKİYE

Received Bachelor's degree in Systems Engineering from the King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia in September 1989.

Completed Master's degree requirements at King Fahd University of Petroleum and Minerals in April 1993.